(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 July 2002 (04.07.2002)

PCT

(10) International Publication Number
WO 02/051231 A1

(51) International Patent Classification: Not classified

(21) International Application Number: PCT/EP00/13143

(22) International Filing Date:
22 December 2000 (22.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **BASF AKTIENGESELLSCHAFT** [DE/DE]; 67056 Ludwigshafen (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **POMPEJUS, Markus** [DE/DE]; Wenjenstr. 21, 67251 Freinsheim (DE). **KRÖGER, Burkhard** [DE/DE]; Im Waldhof 1, 67117 Limburgerhof (DE). **ZELDER, Oskar** [DE/DE]; Rossmarktstr. 27, 67346 Speyer (DE). **SCHRÖDER, Hartwig** [DE/DE]; Benzstr. 4, 59226 Nussloch (DE).

(74) Common Representative: **BASF AKTIENGESELLSCHAFT**; 67056 Ludwigshafen (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/051231 A1

(54) Title: GENES OF CORYNEBACTERIUM

(57) Abstract: Isolated nucleic acid molecules, designated MP nucleic acid molecules, which encode novel MP proteins from Corynebacterium glutamicum are described. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing MP nucleic acid molecules, and host cells into which the expression vectors have been introduced. The invention still further provides isolated MP proteins, mutated MP proteins, fusion proteins, antigenic peptides and methods for the improvement of production of a desired compound from C. glutamicum based on genetic engineering of MP genes in this organism.

GENES OF CORYNEBACTERIUM

Isolated nucleic acid molecules, designated MP nucleic acid molecules, which encode novel MP proteins from *Corynebacterium glutamicum* are described. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing MP nucleic acid molecules, and host cells into which the expression vectors have been introduced. The invention still further provides isolated MP proteins, mutated MP proteins, fusion proteins, antigenic peptides and methods for the improvement of production of a desired compound from *C. glutamicum* based on genetic engineering of MP genes in this organism.

Certain products and by-products of naturally-occurring metabolic processes in cells have utility in a wide array of industries, including the food, feed, cosmetics, and pharmaceutical industries. These molecules, collectively termed 'fine chemicals', include organic acids, both proteinogenic and non-proteinogenic amino acids, nucleotides and nucleosides, lipids and fatty acids, diols, carbohydrates, aromatic compounds, vitamins and cofactors, and enzymes. Their production is most conveniently performed through large-scale culture of bacteria developed to produce and secrete large quantities of a particular desired molecule. One particularly useful organism for this purpose is *Corynebacterium glutamicum*, a gram positive, nonpathogenic bacterium. Through strain selection, a number of mutant strains have been developed which produce an array of desirable compounds. However, selection of strains improved for the production of a particular molecule is a time-consuming and difficult process.

The invention provides novel bacterial nucleic acid molecules which have a variety of uses. These uses include the identification of microorganisms which can be used to produce fine chemicals, the modulation of fine chemical production in *C. glutamicum* or related bacteria, the typing or identification of *C. glutamicum* or related bacteria, as reference points for mapping the *C. glutamicum* genome, and as markers for transformation. These novel nucleic acid molecules encode proteins, referred to herein as metabolic pathway (MP) proteins.

C. glutamicum is a gram positive, aerobic bacterium which is commonly used in industry for the large-scale production of a variety of fine chemicals, and also for the degradation of hydrocarbons (such as in petroleum spills) and for the oxidation

of terpenoids. The MP nucleic acid molecules of the invention, therefore, can be used to identify microorganisms which can be used to produce fine chemicals, e.g., by fermentation processes. Modulation of the expression of the MP nucleic acids of the invention, or modification of the sequence of the MP nucleic acid molecules of the invention, can be used to modulate the production of one or more fine chemicals from a microorganism (e.g., to improve the yield or production of one or more fine chemicals from a Corynebacterium or Brevibacterium species).

The MP nucleic acids of the invention may also be used to identify an organism as being *Corynebacterium glutamicum* or a close relative thereof, or to identify the presence of *C. glutamicum* or a relative thereof in a mixed population of microorganisms. The invention provides the nucleic acid sequences of a number of *C. glutamicum* genes; by probing the extracted genomic DNA of a culture of a unique or mixed population of microorganisms under stringent conditions with a probe spanning a region of a *C. glutamicum* gene which is unique to this organism, one can ascertain whether this organism is present. Although *Corynebacterium glutamicum* itself is nonpathogenic, it is related to species pathogenic in humans, such as *Corynebacterium diphtheriae* (the causative agent of diphtheria); the detection of such organisms is of significant clinical relevance.

The MP nucleic acid molecules of the invention may also serve as reference points for mapping of the *C. glutamicum* genome, or of genomes of related organisms. Similarly, these molecules, or variants or portions thereof, may serve as markers for genetically engineered Corynebacterium or Brevibacterium species.

The MP proteins encoded by the novel nucleic acid molecules of the invention are capable of, for example, performing an enzymatic step involved in the metabolism of certain fine chemicals, including amino acids, vitamins, cofactors, nutraceuticals, nucleotides, nucleosides, and trehalose. Given the availability of cloning vectors for use in *Corynebacterium glutamicum*, such as those disclosed in Sinskey *et al.*, U.S. Patent No. 4,649,119, and techniques for genetic manipulation of *C. glutamicum* and the related *Brevibacterium* species (*e.g., lactofermentum*) (Yoshihama et al, *J. Bacteriol.* 162: 591-597 (1985); Katsumata *et al.*, *J. Bacteriol.* 159: 306-311 (1984); and Santamaria *et al.*, *J. Gen. Microbiol.* 130: 2237-2246 (1984)), the nucleic acid molecules of the invention may be utilized in the genetic engineering of this organism to make it a better or more efficient producer of one or more fine chemicals.

This improved production or efficiency of production of a fine chemical may be due to a direct effect of manipulation of a gene of the invention, or it may be due to an indirect effect of such manipulation. Specifically, alterations in *C. glutamicum* metabolic pathways for amino acids, vitamins, cofactors, nucleotides, and trehalose may have a direct impact on the overall production of one or more of these desired compounds from this organism. For example, optimizing the activity of a trehalose or a lysine or a methionine biosynthetic pathway protein or decreasing the activity of a trehalose or a lysine or methionine degradative pathway protein may result in an increase in the yield or efficiency of production of trehalose or lysine or methionine from such an engineered organism. Alterations in the proteins involved in these metabolic pathways may also have an indirect impact on the production or efficiency of production of a desired fine chemical. For example, a reaction which is in competition for an intermediate necessary for the production of a desired molecule may be eliminated, or a pathway necessary for the production of a particular intermediate for a desired compound may be optimized. Further, modulations in the biosynthesis or degradation of, for example, an amino acid, a vitamin, or a nucleotide may increase the overall ability of the microorganism to rapidly grow and divide, thus increasing the number and/or production capacities of the microorganism in culture and thereby increasing the possible yield of the desired fine chemical.

The nucleic acid and protein molecules of the invention may be utilized to directly improve the production or efficiency of production of one or more desired fine chemicals from *Corynebacterium glutamicum*. Using recombinant genetic techniques well known in the art, one or more of the biosynthetic or degradative enzymes of the invention for amino acids, vitamins, cofactors, nutraceuticals, nucleotides, nucleosides, or trehalose may be manipulated such that its function is modulated. For example, a biosynthetic enzyme may be improved in efficiency, or its allosteric control region destroyed such that feedback inhibition of production of the compound is prevented. Similarly, a degradative enzyme may be deleted or modified by substitution, deletion, or addition such that its degradative activity is lessened for the desired compound without impairing the viability of the cell. In each case, the overall yield or rate of production of the desired fine chemical may be increased.

It is also possible that such alterations in the protein and nucleotide molecules of the invention may improve the production of other fine chemicals besides the amino acids, vitamins,

cofactors, nutraceuticals, nucleotides, nucleosides, and trehalose through indirect mechanisms. Metabolism of any one compound is necessarily intertwined with other biosynthetic and degradative pathways within the cell, and necessary cofactors, intermediates, or substrates in one pathway are likely supplied or limited by another such pathway. Therefore, by modulating the activity of one or more of the proteins of the invention, the production or efficiency of activity of another fine chemical biosynthetic or degradative pathway may be impacted. For example, amino acids serve as the structural units of all proteins, yet may be present intracellularly in levels which are limiting for protein synthesis; therefore, by increasing the efficiency of production or the yields of one or more amino acids within the cell, proteins, such as biosynthetic or degradative proteins, may be more readily synthesized. Likewise, an alteration in a metabolic pathway enzyme such that a particular side reaction becomes more or less favored may result in the over- or under-production of one or more compounds which are utilized as intermediates or substrates for the production of a desired fine chemical.

This invention provides novel nucleic acid molecules which encode proteins, referred to herein as metabolic pathway proteins (MP), which are capable of, for example, performing an enzymatic step involved in the metabolism of molecules important for the normal functioning of cells, such as amino acids, vitamins, cofactors, nucleotides and nucleosides, or trehalose. Nucleic acid molecules encoding an MP protein are referred to herein as MP nucleic acid molecules. In a preferred embodiment, the MP protein performs an enzymatic step related to the metabolism of one or more of the following: amino acids, vitamins, cofactors, nutraceuticals, nucleotides, nucleosides, and trehalose. Examples of such proteins include those encoded by the genes set forth in Table 1.

Table 1: Genes in the Application

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Gene name (identifier)	Function
1	2	metH	5--Methyltetrahydrofolate--homocysteine methyltransferase (EC 2.1.1.13)
3	4	treS	Trehalose Synthase

Accordingly, one aspect of the invention pertains to isolated nucleic acid molecules (*e.g.*, cDNAs, DNAs, or RNAs) comprising a nucleotide sequence encoding an MP protein or biologically active portions thereof, as well as nucleic acid fragments suitable as

primers or hybridization probes for the detection or amplification of MP-encoding nucleic acid (*e.g.*, DNA or mRNA). In particularly preferred embodiments, the isolated nucleic acid molecule comprises one of the nucleotide sequences set forth as the odd-numbered SEQ ID NOs in the Sequence Listing (SEQ ID NO:1, SEQ ID NO:3), or the coding region or a complement thereof of one of these nucleotide sequences. In other particularly preferred embodiments, the isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes to or is at least about 63%, preferably at least about 71%, more preferably at least about 75%, 80% or 90%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence which encodes a proteine sequence set forth as an even-numbered SEQ ID NO in the Sequence Listing (SEQ ID NO:2, SEQ ID NO:4), or a portion thereof. In other preferred embodiments, the isolated nucleic acid molecule encodes one of the amino acid sequences set forth as an even-numbered SEQ ID NO in the Sequence Listing (SEQ ID NO:2, SEQ ID NO:4). The preferred MP proteins of the present invention also preferably possess at least one of the MP activities described herein.

In another embodiment, the isolated nucleic acid molecule encodes a protein or portion thereof wherein the protein or portion thereof includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of the invention (*e.g.*, a sequence having an even-numbered SEQ ID NO: in the Sequence Listing), *e.g.*, sufficiently homologous to an amino acid sequence of the invention such that the protein or portion thereof maintains an MP activity. Preferably, the protein or portion thereof encoded by the nucleic acid molecule maintains the ability to perform an enzymatic reaction in a amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathway. In one embodiment, the protein encoded by the nucleic acid molecule is at least about 63%, preferably at least about 71%, and more preferably at least about 75%, 80%, or 90% and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an amino acid sequence of the invention (*e.g.*, an entire amino acid sequence selected from those having an even-numbered SEQ ID NO in the Sequence Listing). In another preferred embodiment, the protein is a full length *C. glutamicum* protein which is substantially homologous to an entire amino acid sequence of the invention (encoded by an open reading frame shown in the corresponding odd-numbered SEQ ID NOs in the Sequence Listing (SEQ ID NO:2, SEQ ID NO:4).

In another preferred embodiment, the isolated nucleic acid molecule is derived from *C. glutamicum* and encodes a protein (*e.g.*, an MP fusion protein) which includes a biologically active domain which is at least about 50% or more homologous to one of the amino acid sequences of the invention (*e.g.*, a sequence of one of the even-numbered SEQ ID NOs in the Sequence Listing) and is able to catalyze a reaction in a metabolic pathway for an amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose, or one or more of the activities set forth in Table 1, and which also includes heterologous nucleic acid sequences encoding a heterologous polypeptide or regulatory regions.

In another embodiment, the isolated nucleic acid molecule is at least 15 nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising a nucleotide sequence of the invention (*e.g.*, a sequence of an odd-numbered SEQ ID NO in the Sequence Listing). Preferably, the isolated nucleic acid molecule corresponds to a naturally-occurring nucleic acid molecule. More preferably, the isolated nucleic acid encodes a naturally-occurring *C. glutamicum* MP protein, or a biologically active portion thereof.

Another aspect of the invention pertains to vectors, *e.g.*, recombinant expression vectors, containing the nucleic acid molecules of the invention, and host cells into which such vectors have been introduced. In one embodiment, such a host cell is used to produce an MP protein by culturing the host cell in a suitable medium. The MP protein can be then isolated from the medium or the host cell.

Yet another aspect of the invention pertains to a genetically altered microorganism in which an MP gene has been introduced or altered. In one embodiment, the genome of the microorganism has been altered by introduction of a nucleic acid molecule of the invention encoding wild-type or mutated MP sequence as a transgene. In another embodiment, an endogenous MP gene within the genome of the microorganism has been altered, *e.g.*, functionally disrupted, by homologous recombination with an altered MP gene. In another embodiment, an endogenous or introduced MP gene in a microorganism has been altered by one or more point mutations, deletions, or inversions, but still encodes a functional MP protein. In still another embodiment, one or more of the regulatory regions (*e.g.*, a promoter, repressor, or inducer) of an MP gene in a microorganism has been altered (*e.g.*, by deletion, truncation, inversion, or point mutation) such that the expression of the MP gene is modulated. In a preferred

embodiment, the microorganism belongs to the genus *Corynebacterium* or *Brevibacterium*, with *Corynebacterium glutamicum* being particularly preferred. In a preferred embodiment, the microorganism is also utilized for the production of a desired compound, such as trehalose or an amino acid, with lysine and methionine being particularly preferred.

In another aspect, the invention provides a method of identifying the presence or activity of *Cornyebacterium diphtheriae* in a subject. This method includes detection of one or more of the nucleic acid or amino acid sequences of the invention (*e.g.*, the sequences set forth in the Sequence Listing as SEQ ID NOs 1 through 4) in a subject, thereby detecting the presence or activity of *Corynebacterium diphtheriae* in the subject.

Still another aspect of the invention pertains to an isolated MP protein or a portion, *e.g.*, a biologically active portion, thereof. In a preferred embodiment, the isolated MP protein or portion thereof can catalyze an enzymatic reaction involved in one or more pathways for the metabolism of an amino acid, a vitamin, a cofactor, a nutraceutical, a nucleotide, a nucleoside, or trehalose. In another preferred embodiment, the isolated MP protein or portion thereof is sufficiently homologous to an amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: in the Sequence Listing) such that the protein or portion thereof maintains the ability to catalyze an enzymatic reaction involved in one or more pathways for the metabolism of an amino acid, a vitamin, a cofactor, a nutraceutical, a nucleotide, a nucleoside, or trehalose.

The invention also provides an isolated preparation of an MP protein. In preferred embodiments, the MP protein comprises an amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing). In another preferred embodiment, the invention pertains to an isolated full length protein which is substantially homologous to an entire amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing) (encoded by an open reading frame set forth in a corresponding odd-numbered SEQ ID NO: of the Sequence Listing). In yet another embodiment, the protein is at least about 63%, preferably at least about 71%, and more preferably at least about 75%, 80%, or 90%, and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an entire amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing). In other embodiments, the isolated MP protein comprises an amino acid sequence which is at least about 63% or more

homologous to one of the amino acid sequences of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing) and is able to catalyze an enzymatic reaction in an amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathway, or has one or more of the activities set forth in Table 1.

Alternatively, the isolated MP protein can comprise an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, or is at least about 63%, preferably at least about 71%, more preferably at least about 70%, 80%, or 90%, and even more preferably at least about 95%, 96%, 97%, 98,%, or 99% or more homologous to a nucleotide sequence encoding a proteine of one of the even-numbered SEQ ID NOs set forth in the Sequence Listing. It is also preferred that the preferred forms of MP proteins also have one or more of the MP bioactivities described herein.

The MP polypeptide, or a biologically active portion thereof, can be operatively linked to a non-MP polypeptide to form a fusion protein. In preferred embodiments, this fusion protein has an activity which differs from that of the MP protein alone. In other preferred embodiments, this fusion protein, when introduced into a *C. glutamicum* pathway for the metabolism of an amino acid, vitamin, cofactor, nutraceutical, results in increased yields and/or efficiency of production of a desired fine chemical from *C. glutamicum*. In particularly preferred embodiments, integration of this fusion protein into an amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathway of a host cell modulates production of a desired compound from the cell.

In another aspect, the invention provides methods for screening molecules which modulate the activity of an MP protein, either by interacting with the protein itself or a substrate or binding partner of the MP protein, or by modulating the transcription or translation of an MP nucleic acid molecule of the invention.

Another aspect of the invention pertains to a method for producing a fine chemical. This method involves the culturing of a cell containing a vector directing the expression of an MP nucleic acid molecule of the invention, such that a fine chemical is produced. In a preferred embodiment, this method further includes the step of obtaining a cell containing such a vector, in which a cell is transfected with a vector directing the expression of an MP nucleic acid. In another preferred embodiment, this method further includes the step of recovering

the fine chemical from the culture. In a particularly preferred embodiment, the cell is from the genus *Corynebacterium* or *Brevibacterium*, or is selected from those strains set forth in Table 2.

Table 2: Corynebacterium and Brevibacterium strains which may be used in the practice of the invention

Genus	Species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NCTC	DSMZ
Brevibacterium	Ammoniagenes	21054							
Brevibacterium	Ammoniagenes	19350							
Brevibacterium	Ammoniagenes	19351							
Brevibacterium	Ammoniagenes	19352							
Brevibacterium	Ammoniagenes	19353							
Brevibacterium	Ammoniagenes	19354							
Brevibacterium	Ammoniagenes	19355							
Brevibacterium	Ammoniagenes	19356							
Brevibacterium	Ammoniagenes	21055							
Brevibacterium	Ammoniagenes	21077							
Brevibacterium	Ammoniagenes	21553							
Brevibacterium	Ammoniagenes	21580							
Brevibacterium	Ammoniagenes	39101							
Brevibacterium	Butanicum	21196							
Brevibacterium	Divaricatum	21792	P928						
Brevibacterium	Flavum	21474							
Brevibacterium	Flavum	21129							
Brevibacterium	Flavum	21518							
Brevibacterium	Flavum			B11474					
Brevibacterium	Flavum			B11472					
Brevibacterium	Flavum	21127							
Brevibacterium	Flavum	21128							
Brevibacterium	Flavum	21427							
Brevibacterium	Flavum	21475							
Brevibacterium	Flavum	21517							
Brevibacterium	Flavum	21528							
Brevibacterium	Flavum	21529							
Brevibacterium	Flavum			B11477					
Brevibacterium	Flavum			B11478					
Brevibacterium	Flavum	21127							
Brevibacterium	Flavum			B11474					
Brevibacterium	Healii	15527							
Brevibacterium	Ketoglutamicum	21004							
Brevibacterium	Ketoglutamicum	21089							
Brevibacterium	Ketosoreductum	21914							
Brevibacterium	Lactofermentum				70				
Brevibacterium	Lactofermentum				74				
Brevibacterium	Lactofermentum				77				
Brevibacterium	Lactofermentum	21798							
Brevibacterium	Lactofermentum	21799							
Brevibacterium	Lactofermentum	21800							
Brevibacterium	Lactofermentum	21801							
Brevibacterium	Lactofermentum			B11470					
Brevibacterium	Lactofermentum			B11471					
Brevibacterium	Lactofermentum	21086							

Genus	Species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NCTC	DSMZ
Brevibacterium	Lactofermentum	21420							
Brevibacterium	Lactofermentum	21086							
Brevibacterium	Lactofermentum	31269							
Brevibacterium	Linens	9174							
Brevibacterium	Linens	19391							
Brevibacterium	Linens	8377							
Brevibacterium	Paraffinolyticum					11160			
Brevibacterium	spec.						717.73		
Brevibacterium	spec.						717.73		
Brevibacterium	spec.	14604							
Brevibacterium	spec.	21860							
Brevibacterium	spec.	21864							
Brevibacterium	spec.	21865							
Brevibacterium	spec.	21866							
Brevibacterium	spec.	19240							
Corynebacterium	Acetoacidophilum	21476							
Corynebacterium	Acetoacidophilum	13870							
Corynebacterium	Acetoglutamicum			B11473					
Corynebacterium	Acetoglutamicum			B11475					
Corynebacterium	Acetoglutamicum	15806							
Corynebacterium	Acetoglutamicum	21491							
Corynebacterium	Acetoglutamicum	31270							
Corynebacterium	Acetophilum			B3671					
Corynebacterium	Ammoniagenes	6872						2399	
Corynebacterium	Ammoniagenes	15511							
Corynebacterium	Fujiokense	21496							
Corynebacterium	Glutamicum	14067							
Corynebacterium	Glutamicum	39137							
Corynebacterium	Glutamicum	21254							
Corynebacterium	Glutamicum	21255							
Corynebacterium	Glutamicum	31830							
Corynebacterium	Glutamicum	13032							
Corynebacterium	Glutamicum	14305							
Corynebacterium	Glutamicum	15455							
Corynebacterium	Glutamicum	13058							

Genus	Species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NCTC	DSMZ
Coryne-bacterium	Glutamicum	13059							
Coryne-bacterium	Glutamicum	13060							
Coryne-bacterium	Glutamicum	21492							
Coryne-bacterium	Glutamicum	21513							
Coryne-bacterium	Glutamicum	21526							
Coryne-bacterium	Glutamicum	21543							
Coryne-bacterium	Glutamicum	13287							
Coryne-bacterium	Glutamicum	21851							
Coryne-bacterium	Glutamicum	21253							
Coryne-bacterium	glutamicum	21514							
Coryne-bacterium	glutamicum	21516							
Coryne-bacterium	glutamicum	21299							
Coryne-bacterium	glutamicum	21300							
Coryne-bacterium	glutamicum	39684							
Coryne-bacterium	glutamicum	21488							
Coryne-bacterium	glutamicum	21649							
Coryne-bacterium	glutamicum	21650							
Coryne-bacterium	glutamicum	19223							
Coryne-bacterium	glutamicum	13869							
Coryne-bacterium	glutamicum	21157							
Coryne-bacterium	glutamicum	21158							
Coryne-bacterium	glutamicum	21159							
Coryne-bacterium	glutamicum	21355							
Coryne-bacterium	glutamicum	31808							
Coryne-bacterium	glutamicum	21674							
Coryne-bacterium	glutamicum	21562							
Coryne-bacterium	glutamicum	21563							
Coryne-bacterium	glutamicum	21564							

Genus	Species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NCTC	DSMZ
Corynebacterium	glutamicum	21565							
Corynebacterium	glutamicum	21566							
Corynebacterium	glutamicum	21567							
Corynebacterium	glutamicum	21568							
Corynebacterium	glutamicum	21569							
Corynebacterium	glutamicum	21570							
Corynebacterium	glutamicum	21571							
Corynebacterium	glutamicum	21572							
Corynebacterium	glutamicum	21573							
Corynebacterium	glutamicum	21579							
Corynebacterium	glutamicum	19049							
Corynebacterium	glutamicum	19050							
Corynebacterium	glutamicum	19051							
Corynebacterium	glutamicum	19052							
Corynebacterium	glutamicum	19053							
Corynebacterium	glutamicum	19054							
Corynebacterium	glutamicum	19055							
Corynebacterium	glutamicum	19056							
Corynebacterium	glutamicum	19057							
Corynebacterium	glutamicum	19058							
Corynebacterium	glutamicum	19059							
Corynebacterium	glutamicum	19060							
Corynebacterium	glutamicum	19185							
Corynebacterium	glutamicum	13286							
Corynebacterium	glutamicum	21515							
Corynebacterium	glutamicum	21527							
Corynebacterium	glutamicum	21544							
Corynebacterium	glutamicum	21492							

Genus	Species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NCTC	DSMZ
Coryne-bacterium	glutamicum			B8183					
Coryne-bacterium	glutamicum			B8182					
Coryne-bacterium	glutamicum			B12416					
Coryne-bacterium	glutamicum			B12417					
Coryne-bacterium	glutamicum			B12418					
Coryne-bacterium	glutamicum			B11476					
Coryne-bacterium	glutamicum	21608							
Coryne-bacterium	lilium		P973						
Coryne-bacterium	nitrilophilus	21419				11594			
Coryne-bacterium	spec.		P4445						
Coryne-bacterium	spec.		P4446						
Coryne-bacterium	spec.	31088							
Coryne-bacterium	spec.	31089							
Coryne-bacterium	spec.	31090							
Coryne-bacterium	spec.	31090							
Coryne-bacterium	spec.	31090							
Coryne-bacterium	spec.	15954							20145
Coryne-bacterium	spec.	21857							
Coryne-bacterium	spec.	21862							
Coryne-bacterium	spec.	21863							

ATCC: American Type Culture Collection, Rockville, MD, USA
FERM: Fermentation Research Institute, Chiba, Japan
NRRL: ARS Culture Collection, Northern Regional Research Laboratory, Peoria, IL, USA
CECT: Coleccion Espanola de Cultivos Tipo, Valencia, Spain
NCIMB: National Collection of Industrial and Marine Bacteria Ltd., Aberdeen, UK
CBS: Centraalbureau voor Schimmelcultures, Baarn, NL
NCTC: National Collection of Type Cultures, London, UK
DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany

Another aspect of the invention pertains to methods for modulating production of a molecule from a microorganism. Such methods include contacting the cell with an agent which modulates MP protein activity or MP nucleic acid expression such that a cell associated activity is altered relative to this same activity in the absence of the agent. In a preferred embodiment, the cell is modulated for one or more *C. glutamicum* amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathways, such that the yields or rate of production of a desired fine chemical by this microorganism is improved. The agent which modulates MP protein activity can be an agent which stimulates MP protein activity or MP nucleic acid expression. Examples of agents which stimulate MP protein activity or MP nucleic acid expression include small molecules, active MP proteins, and nucleic acids encoding MP proteins that have been introduced into the cell. Examples of agents which inhibit MP activity or expression include small molecules, and antisense MP nucleic acid molecules.

Another aspect of the invention pertains to methods for modulating yields of a desired compound from a cell, involving the introduction of a wild-type or mutant MP gene into a cell, either maintained on a separate plasmid or integrated into the genome of the host cell. If integrated into the genome, such integration can be random, or it can take place by homologous recombination such that the native gene is replaced by the introduced copy, causing the production of the desired compound from the cell to be modulated. In a preferred embodiment, said yields are increased. In another preferred embodiment, said chemical is a fine chemical. In a particularly preferred embodiment, said fine chemical is trehalose or an amino acid. In especially preferred embodiments, said amino acid are L-lysine and L-methionine.

Detailed Description of the Invention

The present invention provides MP nucleic acid and protein molecules which are involved in the metabolism of certain fine chemicals in *Corynebacterium glutamicum*, including amino acids, vitamins, cofactors, nutraceuticals, nucleotides, nucleosides, and trehalose. The molecules of the invention may be utilized in the modulation of production of fine chemicals from microorganisms, such as *C. glutamicum*, either directly (*e.g.*, where modulation of the activity of a trehalose or a lysine or methionine biosynthesis protein has a direct impact on the production or efficiency of production of trehalose or lysine or methionine from that organism), or may have an indirect impact

which nonetheless results in an increase of yield or efficiency of production of the desired compound (*e.g.*, where modulation of the activity of a nucleotide biosynthesis protein has an impact on the production of an organic acid or a fatty acid from the bacterium, perhaps due to improved growth or an increased supply of necessary co-factors, energy compounds, or precursor molecules). Aspects of the invention are further explicated below.

1. Fine Chemicals

The term 'fine chemical' is art-recognized and includes molecules produced by an organism which have applications in various industries, such as, but not limited to, the pharmaceutical, agriculture, and cosmetics industries. Such compounds include organic acids, such as tartaric acid, itaconic acid, and diaminopimelic acid, both proteinogenic and non-proteinogenic amino acids, purine and pyrimidine bases, nucleosides, and nucleotides (as described *e.g.* in Kuninaka, A. (1996) Nucleotides and related compounds, p. 561-612, in Biotechnology vol. 6, Rehm *et al.*, eds. VCH: Weinheim, and references contained therein), lipids, both saturated and unsaturated fatty acids (*e.g.*, arachidonic acid), diols (*e.g.*, propane diol, and butane diol), carbohydrates (*e.g.*, hyaluronic acid and trehalose), aromatic compounds (*e.g.*, aromatic amines, vanillin, and indigo), vitamins and cofactors (as described in Ullmann's Encyclopedia of Industrial Chemistry, vol. A27, "Vitamins", p. 443-613 (1996) VCH: Weinheim and references therein; and Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research - Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press, (1995)), enzymes, polyketides (Cane *et al.* (1998) *Science* 282: 63-68), and all other chemicals described in Gutcho (1983) Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 and references therein. The metabolism and uses of certain of these fine chemicals are further explicated below.

A. *Amino Acid Metabolism and Uses*

Amino acids comprise the basic structural units of all proteins, and as such are essential for normal cellular functioning in all organisms. The term "amino acid" is art-recognized. The proteinogenic amino acids, of which there are 20 species, serve as structural units for proteins, in which they are linked by peptide bonds, while the nonproteinogenic amino acids (hundreds of which are known) are not normally found in proteins (see

Ulmann's Encyclopedia of Industrial Chemistry, vol. A2, p. 57-97 VCH: Weinheim (1985)). Amino acids may be in the D- or L- optical configuration, though L-amino acids are generally the only type found in naturally-occurring proteins. Biosynthetic and degradative pathways of each of the 20 proteinogenic amino acids have been well characterized in both prokaryotic and eukaryotic cells (see, for example, Stryer, L. Biochemistry, 3rd edition, pages 578-590 (1988)). The 'essential' amino acids (histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine), so named because they are generally a nutritional requirement due to the complexity of their biosyntheses, are readily converted by simple biosynthetic pathways to the remaining 11 'nonessential' amino acids (alanine, arginine, asparagine, aspartate, cysteine, glutamate, glutamine, glycine, proline, serine, and tyrosine). Higher animals do retain the ability to synthesize some of these amino acids, but the essential amino acids must be supplied from the diet in order for normal protein synthesis to occur.

Aside from their function in protein biosynthesis, these amino acids are interesting chemicals in their own right, and many have been found to have various applications in the food, feed, chemical, cosmetics, agriculture, and pharmaceutical industries. Lysine is an important amino acid in the nutrition not only of humans, but also of monogastric animals such as poultry and swine. Glutamate is most commonly used as a flavor additive (mono-sodium glutamate, MSG) and is widely used throughout the food industry, as are aspartate, phenylalanine, glycine, and cysteine. Glycine, L-methionine and tryptophan are all utilized in the pharmaceutical industry. Glutamine, valine, leucine, isoleucine, histidine, arginine, proline, serine and alanine are of use in both the pharmaceutical and cosmetics industries. Threonine, tryptophan, and D/ L-methionine are common feed additives. (Leuchtenberger, W. (1996) Amino aids - technical production and use, p. 466-502 in Rehm *et al.* (eds.) Biotechnology vol. 6, chapter 14a, VCH: Weinheim). Additionally, these amino acids have been found to be useful as precursors for the synthesis of synthetic amino acids and proteins, such as N-acetylcysteine, S-carboxymethyl-L-cysteine, (S)-5-hydroxytryptophan, and others described in Ulmann's Encyclopedia of Industrial Chemistry, vol. A2, p. 57-97, VCH: Weinheim, 1985.

The biosynthesis of these natural amino acids in organisms capable of producing them, such as bacteria, has been well characterized (for review of bacterial amino acid biosynthesis and regulation thereof, see Umbarger, H.E.(1978) *Ann. Rev.*

Biochem. 47: 533-606). Glutamate is synthesized by the reductive amination of á-ketoglutarate, an intermediate in the citric acid cycle. Glutamine, proline, and arginine are each subsequently produced from glutamate. The biosynthesis of serine is a three-step process beginning with 3-phosphoglycerate (an intermediate in glycolysis), and resulting in this amino acid after oxidation, transamination, and hydrolysis steps. Both cysteine and glycine are produced from serine; the former by the condensation of homocysteine with serine, and the latter by the transferal of the side-chain â-carbon atom to tetrahydrofolate, in a reaction catalyzed by serine transhydroxymethylase. Phenylalanine, and tyrosine are synthesized from the glycolytic and pentose phosphate pathway precursors erythrose 4-phosphate and phosphoenolpyruvate in a 9-step biosynthetic pathway that differ only at the final two steps after synthesis of prephenate. Tryptophan is also produced from these two initial molecules, but its synthesis is an 11-step pathway. Tyrosine may also be synthesized from phenylalanine, in a reaction catalyzed by phenylalanine hydroxylase. Alanine, valine, and leucine are all biosynthetic products of pyruvate, the final product of glycolysis. Aspartate is formed from oxaloacetate, an intermediate of the citric acid cycle. Asparagine, methionine, threonine, and lysine are each produced by the conversion of aspartate. Isoleucine is formed from threonine. The biosynthetic pathways leading to methionine have been studied in diverse organisms and show similarity as well as differences. The first step, acylation of homoserine, is common to all the organisms, even though the source of the transferred acyl groups is different. *Escherichia coli* and the related species use succinyl-CoA (Michaeli, S. and Ron, E. Z. (1981) Construction and physical mapping of plasmids containing the *metA* gene of *Escherichia coli* K12, *Mol. Gen. Genet.* **182**, 349-354). Construction and physical mapping of plasmids containing the *metA* gene of *Escherichia coli* K12, *Mol. Gen. Genet.* **182**, 349-354), while *Saccharomyces cerevisiae* (Langin, T., Faugeron, G., Goyon, C., Nicolas, A., and Rossignol, J. (1986) The MET2 gene of *Saccharomyces cerevisiae*: molecular cloning and nucleotide sequence. *Gene* **49**, 283-293), *Brevibacterium flavum* (Miyajima, R. and Shiio, I. (1973) Regulation of aspartate family of amino acid biosynthesis in *Brevibacterium flavum*: properties of homoserine *O*-transacetylase. *J. Biochem.* **73**, 1061-1068; Ozaki, H. and Shiio, I. (1982) Methionine biosynthesis in *Brevibacterium flavum*: properties and essential role of *O*-acetylhomoserine sulfhydrylase. *J. Biochem.* **91**, 1163-1171), *C. glutamicum* (Park, S.-D., Lee, J.-Y., Kim, Y., Kim, J.-H., and Lee, H.-S. (1998) Isolation and analysis of *metA*, a methionine biosynthetic gene encoding homoserine acetyltransferase in *Corynebacterium*

glutamicum. *Mol. Cells* **8**, 286-294), and *Leptospira meyeri* (Belfaiza, J., Martel, A., Maegarita. D., and Saint Girons, I. (1998) Direct sulfhydrylation for methionine biosynthesis in *Leptospira meyeri*. *J. Bacteriol.* **180**, 250-255; Bourhy, P., Martel, A., Margarita, D., Saint Girons, I., and Belfaiza, J. (1997) Homoserine *O*-acetyltransferase, involved in the *Leptospira meyeri* methionine biosynthetic pathway, is not feedback inhibited. *J. Bacteriol.* **179**, 4396-4398) use acetyl-CoA as the acyl donor. Formation of homocysteine from acylhomoserine can occur in two different ways. *E. coli* uses the transsulfuration pathway which is catalyzed by cystathionine γ-synthase (the product of *metB*) and cystathionine β-lyase (the product of *metC*). *S. cerevisiae* (Cherest, H. and Surdin-Kerjan, Y. (1992) Genetic analysis of a new mutation conferring cysteine auxotrophy in *Saccharomyces cerevisiae*: updating of the sulfur metabolism pathway. *Genetics* **130**, 51-58), *B. flavum* (Ozaki, H. and Shiio, I. (1982) Methionine biosynthesis in *Brevibacterium flavum*: properties and essential role of *O*-acetylhomoserine sulfhydrylase. *J. Biochem.* **91**, 1163-1171), *Pseudomonas aeruginosa* (Foglino, M., Borne, F., Bally, M., Ball, G., and Patte, J. C. (1995) A direct sulfhydrylation pathway is used for methionine biosynthesis in *Pseudomonas aeruginosa*. *Microbiology* **141**, 431-439), and *L. meyeri* (Belfaiza, J., Martel, A., Maegarita. D., and Saint Girons, I. (1998) Direct sulfhydrylation for methionine biosynthesis in *Leptospira meyeri*. *J. Bacteriol.* **180**, 250-255) utilize the direct sulfhydrylation pathway which is catalyzed by acylhomoserine sulfhydrylase. Unlike closely related *B. flavum* which uses only the direct sulfhydrylation pathway, enzyme activities of the transsulfuration pathway have been detected in the extracts of the *C. glutamicum* cells and the pathway has been proposed to be the route for methionine biosynthesis in the organism (Hwang, B-J., Kim, Y., Kim, H.-B., Kim, J., Hwang, H.-J., and Lee, H.-S. (1999) Analysis of *Corynebacterium glutamicum* methionine biosynthetic pathway: Isolation and analysis of *metB* encoding cystathionine ã-synthase. *Mol. Cells* **9**, 300-308; Kase, H. and Nakayama, K. (1974) Production of *O*-acetyl-L-homoserine by methionine analog resistant mutants and regulation of homoserine-*O*-transacetylase in *Corynebacterium glutamicum*. *Agr. Biol. Chem.* **38**, 2021-2030; Park, S.-D., Lee, J.-Y., Kim, Y., Kim, J.-H., and Lee, H.-S. (1998) Isolation and analysis of *metA*, a methionine biosynthetic gene encoding homoserine acetyltransferase in *Corynebacterium glutamicum*. *Mol. Cells* **8**, 286-294).

Even though some genes involved in methionine biosynthesis in *C. glutamicum* were isolated in recent years, the information on the biosynthesis of methionine in *C. glutamicum* is still limited.

The metA and metB genes have been isolated from the organism and also the metC and the metZ gene are known (table 4), but the final step of the biosynthesis remained unclear. In this invention, the biosynthetic pathway leading to methionine in *C. glutamicum* is deciphered in total and the biosynthetic gene responsible for the last step of the biosynthesis is defined with the metH gene encoding the enzyme methionine synthase.

A complex 9-step pathway results in the production of histidine from 5-phosphoribosyl-1-pyrophosphate, an activated sugar.

Amino acids in excess of the protein synthesis needs of the cell cannot be stored, and are instead degraded to provide intermediates for the major metabolic pathways of the cell (for review see Stryer, L. Biochemistry 3rd ed. Ch. 21 "Amino Acid Degradation and the Urea Cycle" p. 495-516 (1988)). Although the cell is able to convert unwanted amino acids into useful metabolic intermediates, amino acid production is costly in terms of energy, precursor molecules, and the enzymes necessary to synthesize them. Thus it is not surprising that amino acid biosynthesis is regulated by feedback inhibition, in which the presence of a particular amino acid serves to slow or entirely stop its own production (for overview of feedback mechanisms in amino acid biosynthetic pathways, see Stryer, L. Biochemistry, 3rd ed. Ch. 24: "Biosynthesis of Amino Acids and Heme" p. 575-600 (1988)). Thus, the output of any particular amino acid is limited by the amount of that amino acid present in the cell.

2.1 Vitamin, Cofactor, and Nutraceutical Metabolism and Uses

Vitamins, cofactors, and nutraceuticals comprise another group of molecules which the higher animals have lost the ability to synthesize and so must ingest, although they are readily synthesized by other organisms, such as bacteria. These molecules are either bioactive substances themselves, or are precursors of biologically active substances which may serve as electron carriers or intermediates in a variety of metabolic pathways. Aside from their nutritive value, these compounds also have significant industrial value as coloring agents, antioxidants, and catalysts or other processing aids. (For an overview of the structure, activity, and industrial applications of these compounds, see, for example, Ullman's Encyclopedia of Industrial Chemistry, "Vitamins" vol. A27, p. 443-613, VCH: Weinheim, 1996.) The term "vitamin" is art-recognized, and includes nutrients which are required by an organism for normal functioning, but which that organism cannot synthesize by itself. The group of vitamins may encompass cofactors and nutraceutical compounds. The

language "cofactor" includes nonproteinaceous compounds required for a normal enzymatic activity to occur. Such compounds may be organic or inorganic; the cofactor molecules of the invention are preferably organic. The term "nutraceutical" includes dietary supplements having health benefits in plants and animals, particularly humans. Examples of such molecules are vitamins, antioxidants, and also certain lipids (*e.g.*, polyunsaturated fatty acids).

The biosynthesis of these molecules in organisms capable of producing them, such as bacteria, has been largely characterized (Ullman's Encyclopedia of Industrial Chemistry, "Vitamins" vol. A27, p. 443-613, VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research - Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press: Champaign, IL X, 374 S).

Thiamin (vitamin B_1) is produced by the chemical coupling of pyrimidine and thiazole moieties. Riboflavin (vitamin B_2) is synthesized from guanosine-5'-triphosphate (GTP) and ribose-5'-phosphate. Riboflavin, in turn, is utilized for the synthesis of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). The family of compounds collectively termed 'vitamin B_6' (*e.g.*, pyridoxine, pyridoxamine, pyridoxa-5'-phosphate, and the commercially used pyridoxin hydrochloride) are all derivatives of the common structural unit, 5-hydroxy-6-methylpyridine. Pantothenate (pantothenic acid, (R)-(+)-N-(2,4-dihydroxy-3,3-dimethyl-1-oxobutyl)-â-alanine) can be produced either by chemical synthesis or by fermentation. The final steps in pantothenate biosynthesis consist of the ATP-driven condensation of â-alanine and pantoic acid. The enzymes responsible for the biosynthesis steps for the conversion to pantoic acid, to â-alanine and for the condensation to panthotenic acid are known. The metabolically active form of pantothenate is Coenzyme A, for which the biosynthesis proceeds in 5 enzymatic steps. Pantothenate, pyridoxal-5'-phosphate, cysteine and ATP are the precursors of Coenzyme A. These enzymes not only catalyze the formation of panthothante, but also the production of (R)-pantoic acid, (R)-pantolacton, (R)-panthenol (provitamin B_5), pantetheine (and its derivatives) and coenzyme A.

Biotin biosynthesis from the precursor molecule pimeloyl-CoA in microorganisms has been studied in detail and several of the genes involved have been identified. Many of the corresponding proteins have been found to also be involved in Fe-cluster synthesis and are members of the nifS class of proteins. Lipoic acid is derived from octanoic acid, and serves as a coenzyme in energy metabolism, where it becomes part of the pyruvate dehydrogenase complex and the á-ketoglutarate dehydrogenase complex. The folates are a group of substances which are all derivatives of folic acid, which is turn is derived from L-glutamic acid, p-amino-benzoic acid and 6-methylpterin. The biosynthesis of folic acid and its derivatives, starting from metabolism intermediates guanosine-5'-triphosphate (GTP), L-glutamic acid and p-amino-benzoic acid has been studied in detail in certain microorganisms.

Corrinoids (such as the cobalamines and particularly vitamin B_{12}) and porphyrines belong to a group of chemicals characterized by a tetrapyrole ring system. The biosynthesis of vitamin B_{12} is sufficiently complex that it has not yet been completely characterized, but many of the enzymes and substrates involved are now known. Nicotinic acid (nicotinate), and nicotinamide are pyridine derivatives which are also termed 'niacin'. Niacin is the precursor of the important coenzymes NAD (nicotinamide adenine dinucleotide) and NADP (nicotinamide adenine dinucleotide phosphate) and their reduced forms.

The large-scale production of these compounds has largely relied on cell-free chemical syntheses, though some of these chemicals have also been produced by large-scale culture of microorganisms, such as riboflavin, Vitamin B_6, pantothenate, and biotin. Only Vitamin B_{12} is produced solely by fermentation, due to the complexity of its synthesis. *In vitro* methodologies require significant inputs of materials and time, often at great cost.

C. Purine, Pyrimidine, Nucleoside and Nucleotide Metabolism and Uses

Purine and pyrimidine metabolism genes and their corresponding proteins are important targets for the therapy of tumor diseases and viral infections. The language "purine" or "pyrimidine" includes the nitrogenous bases which are constituents of nucleic acids, co-enzymes, and nucleotides. The term "nucleotide" includes the basic structural units of nucleic acid molecules, which are comprised of a nitrogenous base, a pentose sugar (in the case of RNA, the sugar is ribose; in the case of DNA, the sugar is D-deoxyribose), and phosphoric acid. The language

"nucleoside" includes molecules which serve as precursors to nucleotides, but which are lacking the phosphoric acid moiety that nucleotides possess. By inhibiting the biosynthesis of these molecules, or their mobilization to form nucleic acid molecules, it is possible to inhibit RNA and DNA synthesis; by inhibiting this activity in a fashion targeted to cancerous cells, the ability of tumor cells to divide and replicate may be inhibited. Additionally, there are nucleotides which do not form nucleic acid molecules, but rather serve as energy stores (*i.e.*, AMP) or as coenzymes (*i.e.*, FAD and NAD).

Several publications have described the use of these chemicals for these medical indications, by influencing purine and/or pyrimidine metabolism (*e.g.* Christopherson, R.I. and Lyons, S.D. (1990) "Potent inhibitors of *de novo* pyrimidine and purine biosynthesis as chemotherapeutic agents." *Med. Res. Reviews* 10: 505-548). Studies of enzymes involved in purine and pyrimidine metabolism have been focused on the development of new drugs which can be used, for example, as immunosuppressants or anti-proliferants (Smith, J.L., (1995) "Enzymes in nucleotide synthesis." *Curr. Opin. Struct. Biol.* 5: 752-757; (1995) *Biochem Soc. Transact.* 23: 877-902). However, purine and pyrimidine bases, nucleosides and nucleotides have other utilities: as intermediates in the biosynthesis of several fine chemicals (*e.g.*, thiamine, *S*-adenosyl-methionine, folates, or riboflavin), as energy carriers for the cell (*e.g.*, ATP or GTP), and for chemicals themselves, commonly used as flavor enhancers (*e.g.*, IMP or GMP) or for several medicinal applications (see, for example, Kuninaka, A. (1996) Nucleotides and Related Compounds in Biotechnology vol. 6, Rehm *et al.*, eds. VCH: Weinheim, p. 561-612). Also, enzymes involved in purine, pyrimidine, nucleoside, or nucleotide metabolism are increasingly serving as targets against which chemicals for crop protection, including fungicides, herbicides and insecticides, are developed.

The metabolism of these compounds in bacteria has been characterized (for reviews see, for example, Zalkin, H. and Dixon, J.E. (1992) "*de novo* purine nucleotide biosynthesis", in: Progress in Nucleic Acid Research and Molecular Biology, vol. 42, Academic Press:, p. 259-287; and Michal, G. (1999) "Nucleotides and Nucleosides", Chapter 8 in: Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Wiley: New York). Purine metabolism has been the subject of intensive research, and is essential to the normal functioning of the cell. Impaired purine metabolism in higher animals can cause severe disease, such as gout. Purine nucleotides are synthesized from ribose-5-phosphate, in a series of steps through the intermediate compound

inosine-5'-phosphate (IMP), resulting in the production of guanosine-5'-monophosphate (GMP) or adenosine-5'-monophosphate (AMP), from which the triphosphate forms utilized as nucleotides are readily formed. These compounds are also utilized as energy stores, so their degradation provides energy for many different biochemical processes in the cell. Pyrimidine biosynthesis proceeds by the formation of uridine-5'-monophosphate (UMP) from ribose-5-phosphate. UMP, in turn, is converted to cytidine-5'-triphosphate (CTP): The deoxy- forms of all of these nucleotides are produced in a one step reduction reaction from the diphosphate ribose form of the nucleotide to the diphosphate deoxyribose form of the nucleotide. Upon phosphorylation, these molecules are able to participate in DNA synthesis.

D. Trehalose Metabolism and Uses

Trehalose consists of two glucose molecules, bound in á, á-1,1 linkage. It is commonly used in the food industry as a sweetener, an additive for dried or frozen foods, and in beverages. However, it also has applications in the pharmaceutical, cosmetics and biotechnology industries (see, for example, Nishimoto *et al.*, (1998) U.S. Patent No. 5,759,610; Singer, M.A. and Lindquist, S. (1998) *Trends Biotech.* 16: 460-467; Paiva, C.L.A. and Panek, A.D. (1996) *Biotech. Ann. Rev.* 2: 293-314; and Shiosaka, M. (1997) J. Japan 172: 97-102). Trehalose is produced by enzymes from many microorganisms and is naturally released into the surrounding medium, from which it can be collected using methods known in the art.

II. Elements and Methods of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as MP nucleic acid and protein molecules, which play a role in or function in one or more cellular metabolic pathways. In one embodiment, the MP molecules catalyze an enzymatic reaction involving one or more amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathways. In a preferred embodiment, the activity of the MP molecules of the present invention in one or more *C. glutamicum* metabolic pathways for amino acids, vitamins, cofactors, nutraceuticals, nucleotides, nucleosides or trehalose has an impact on the production of a desired fine chemical by this organism. In a particularly preferred embodiment, the MP molecules of the invention are modulated in activity, such that the *C. glutamicum* metabolic pathways in which the MP proteins of the invention are involved are modulated in efficiency or output, which either directly or

indirectly modulates the production or efficiency of production of a desired fine chemical by *C. glutamicum*. The MP molecules may be combined with other MP molecules of the same or different metabolic pathway to increase the yield of a desired fine chemical, preferred trehalose or an amino acid, more preferred lysine or methionine. Alternatively or in addition a byproduct which is not desired may be reduced by combination of disruption of MP molecules or other metabolic molecules. The MP molecules combined with other MP molecules of the same or a different pathway may be altered in their nucleotide and in the corresponding amino acid sequence in such a way that their activity is altered under physiological conditions which leads to an increase in productivity and/or yield of a desired fine chemical. In a further embodiment the MP molecule in its original or in its above described altered form may be combined with other MP molecules of the same or a different pathway wich are altered in their nucleotide sequence in such a way that their activity is altered under physiological conditions which leads to an increase in productivity and/or yield of a desired fine chemical. Preferred combinations are such which combine one ore both MP molecules of table1 with one ore more single or multipe copies of MP proteins of tables 4 and 5 or the respective published MP molecules of the same metabolic pathway (Methionine biosyntesis or trehalose/phosphoenolpyruvat way).

The language, "MP protein" or "MP polypeptide" includes proteins which play a role in, *e.g.*, catalyze an enzymatic reaction, in one or more amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside or trehalose metabolic pathways. Examples of MP proteins include those encoded by the MP genes set forth in Table 1 and by the odd-numbered SEQ ID NOs. The terms "MP gene" or "MP nucleic acid sequence" include nucleic acid sequences encoding an MP protein, which consist of a coding region and also corresponding untranslated 5' and 3' sequence regions. Examples of MP genes include those set forth in Table 1. The terms "production" or "productivity" are art-recognized and include the concentration of the fermentation product (for example, the desired fine chemical) formed within a given time and a given fermentation volume (*e.g.*, kg product per hour per liter). The term "efficiency of production" includes the time required for a particular level of production to be achieved (for example, how long it takes for the cell to attain a particular rate of output of a fine chemical). The term "yield" or "product/carbon yield" is art-recognized and includes the efficiency of the conversion of the carbon source into the product (*i.e.*, fine chemical). This is generally written as, for example, kg product per kg carbon source. By increasing the yield or production of the compound,

the quantity of recovered molecules, or of useful recovered molecules of that compound in a given amount of culture over a given amount of time is increased. The terms "biosynthesis" or a "biosynthetic pathway" are art-recognized and include the synthesis of a compound, preferably an organic compound, by a cell from intermediate compounds in what may be a multistep and highly regulated process. The terms "degradation" or a "degradation pathway" are art-recognized and include the breakdown of a compound, preferably an organic compound, by a cell to degradation products (generally speaking, smaller or less complex molecules) in what may be a multistep and highly regulated process. The language "metabolism" is art-recognized and includes the totality of the biochemical reactions that take place in an organism. The metabolism of a particular compound, then, (*e.g.*, the metabolism of an amino acid such as glycine) comprises the overall biosynthetic, modification, and degradation pathways in the cell related to this compound.

In another embodiment, the MP molecules of the invention are capable of modulating the production of a desired molecule, such as a fine chemical, in a microorganism such as *C. glutamicum*. Using recombinant genetic techniques, one or more of the biosynthetic or degradative enzymes of the invention for amino acids, vitamins, cofactors, nutraceuticals, nucleotides, nucleosides, or trehalose may be manipulated such that its function is modulated. For example, a biosynthetic enzyme may be improved in efficiency, or its allosteric control region destroyed such that feedback inhibition of production of the compound is prevented. Similarly, a degradative enzyme may be deleted or modified by substitution, deletion, or addition such that its degradative activity is lessened for the desired compound without impairing the viability of the cell. In each case, the overall yield or rate of production of one of these desired fine chemicals may be increased.

It is also possible that such alterations in the protein and nucleotide molecules of the invention may improve the production of other fine chemicals besides the amino acids, vitamins, cofactors, nutraceuticals, nucleotides, nucleosides, and trehalose. Metabolism of any one compound is necessarily intertwined with other biosynthetic and degradative pathways within the cell, and necessary cofactors, intermediates, or substrates in one pathway are likely supplied or limited by another such pathway. Therefore, by modulating the activity of one or more of the proteins of the invention, the production or efficiency of activity of another fine chemical biosynthetic or degradative pathway may be impacted. For example, amino acids

serve as the structural units of all proteins, yet may be present intracellularly in levels which are limiting for protein synthesis; therefore, by increasing the efficiency of production or the yields of one or more amino acids within the cell, proteins, such as biosynthetic or degradative proteins, may be more readily synthesized. Likewise, an alteration in a metabolic pathway enzyme such that a particular side reaction becomes more or less favored may result in the over- or under-production of one or more compounds which are utilized as intermediates or substrates for the production of a desired fine chemical.

The isolated nucleic acid sequences of the invention are contained within the genome of a *Corynebacterium glutamicum* strain available through the American Type Culture Collection, given designation ATCC 13032. The nucleotide sequence of the isolated *C. glutamicum* MP DNAs and the predicted amino acid sequences of the *C. glutamicum* MP proteins are shown in the Sequence Listing as odd-numbered SEQ ID NOs and even-numbered SEQ ID NOs, respectively. Computational analyses were performed which classified and/or identified these nucleotide sequences as sequences which encode metabolic pathway proteins.

The present invention also pertains to proteins which have an amino acid sequence which is substantially homologous to an amino acid sequence of the invention (*e.g.*, the sequence of an even-numbered SEQ ID NO of the Sequence Listing). As used herein, a protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence is least about 50% homologous to the selected amino acid sequence, *e.g.*, the entire selected amino acid sequence. A protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence can also be least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, or 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to the selected amino acid sequence.

The MP protein or a biologically active portion or fragment thereof of the invention can catalyze an enzymatic reaction in one or more amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathways, or have one or more of the activities set forth in Table 1.

Various aspects of the invention are described in further detail in the following subsections:

A. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode MP polypeptides or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes or primers for the identification or amplification of MP-encoding nucleic acid (*e.g.*, MP DNA). As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. This term also encompasses untranslated sequence located at both the 3' and 5' ends of the coding region of the gene: at least about 100 nucleotides of sequence upstream from the 5' end of the coding region and at least about 20 nucleotides of sequence downstream from the 3'end of the coding region of the gene. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated MP nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (e.g, a *C. glutamicum* cell). Moreover, an "isolated" nucleic acid molecule, such as a DNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having a nucleotide sequence of an odd-numbered SEQ ID NO of the Sequence Listing, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, a *C. glutamicum* MP DNA can be isolated from a *C. glutamicum* library using all or portion of one of the odd-numbered SEQ ID NO sequences of the Sequence Listing as a hybridization probe and standard hybridization techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning:*

A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Moreover, a nucleic acid molecule encompassing all or a portion of one of the nucleic acid sequences of the invention (*e.g.*, an odd-numbered SEQ ID NO:) can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this sequence (*e.g.*, a nucleic acid molecule encompassing all or a portion of one of the nucleic acid sequences of the invention (*e.g.*, an odd-numbered SEQ ID NO of the Sequence Listing) can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this same sequence). For example, mRNA can be isolated from normal endothelial cells (*e.g.*, by the guanidinium-thiocyanate extraction procedure of Chirgwin *et al.* (1979) *Biochemistry* 18: 5294-5299) and DNA can be prepared using reverse transcriptase (*e.g.*, Moloney MLV reverse transcriptase, available from Gibco/BRL, Bethesda, MD; or AMV reverse transcriptase, available from Seikagaku America, Inc., St. Petersburg, FL). Synthetic oligonucleotide primers for polymerase chain reaction amplification can be designed based upon one of the nucleotide sequences shown in the Sequence Listing. A nucleic acid of the invention can be amplified using cDNA or, alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to an MP nucleotide sequence can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises one of the nucleotide sequences shown in the Sequence Listing. The nucleic acid sequences of the invention, as set forth in the Sequence Listing, correspond to the *Corynebacterium glutamicum* MP DNAs of the invention. This DNA comprises sequences encoding MP proteins (*i.e.*, the "coding region", indicated in each odd-numbered SEQ ID NO: sequence in the Sequence Listing), as well as 5' untranslated sequences and 3' untranslated sequences, also indicated in each odd-numbered SEQ ID NO: in the Sequence Listing. Alternatively, the nucleic acid molecule can comprise only the coding region of any of the nucleic acid sequences of the Sequence Listing.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of one of the nucleotide sequences of the invention (*e.g.*, a sequence of an odd-numbered SEQ ID NO: of the

Sequence Listing), or a portion thereof. A nucleic acid molecule which is complementary to one of the nucleotide sequences of the invention is one which is sufficiently complementary to one of the nucleotide sequences shown in the Sequence Listing (*e.g.*, the sequence of an odd-numbered SEQ ID NO:) such that it can hybridize to one of the nucleotide sequences of the invention, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which is at least about 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence of the invention (*e.g.*, a sequence of an odd-numbered SEQ ID NO: of the Sequence Listing), or a portion thereof. Ranges and identity values intermediate to the above-recited ranges, (*e.g.*, 70-90% identical or 80-95% identical) are also intended to be encompassed by the present invention. For example, ranges of identity values using a combination of any of the above values recited as upper and/or lower limits are intended to be included. In an additional preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, to one of the nucleotide sequences of the invention, or a portion thereof.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the coding region of the sequence of one of the odd-numbered SEQ ID NOs of the Sequence Listing, for example a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of an MP protein. The nucleotide sequences determined from the cloning of the MP genes from *C. glutamicum* allows for the generation of probes and primers designed for use in identifying and/or cloning MP homologues in other cell types and organisms, as well as MP homologues from other *Corynebacteria* or related species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 40, 50 or 75 consecutive nucleotides of a sense strand of one of the nucleotide sequences of the invention (*e.g.*, a sequence of one of the odd-numbered SEQ ID NOs of the Sequence Listing), an anti-sense sequence of one of these sequences, or naturally occurring mutants thereof. Primers based on a nucleotide sequence

of the invention can be used in PCR reactions to clone MP homologues. Probes based on the MP nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.* the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells which misexpress an MP protein, such as by measuring a level of an MP-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting MP mRNA levels or determining whether a genomic MP gene has been mutated or deleted.

In one embodiment, the nucleic acid molecule of the invention encodes a protein or portion thereof which includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO of the Sequence Listing) such that the protein or portion thereof maintains the ability to catalyze an enzymatic reaction in an amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathway. As used herein, the language "sufficiently homologous" refers to proteins or portions thereof which have amino acid sequences which include a minimum number of identical or equivalent (*e.g.*, an amino acid residue which has a similar side chain as an amino acid residue in a sequence of one of the even-numbered SEQ ID NOs of the Sequence Listing) amino acid residues to an amino acid sequence of the invention such that the protein or portion thereof is able to catalyze an enzymatic reaction in a *C. glutamicum* amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside or trehalose metabolic pathway. Protein members of such metabolic pathways, as described herein, function to catalyze the biosynthesis or degradation of one or more of: amino acids, vitamins, cofactors, nutraceuticals, nucleotides, nucleosides, or trehalose. Examples of such activities are also described herein. Thus, "the function of an MP protein" contributes to the overall functioning of one or more such metabolic pathway and contributes, either directly or indirectly, to the yield, production, and/or efficiency of production of one or more fine chemicals. Examples of MP protein activities are set forth in Table 1.

In another embodiment, the protein is at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to an entire amino acid

sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing).

Portions of proteins encoded by the MP nucleic acid molecules of the invention are preferably biologically active portions of one of the MP proteins. As used herein, the term "biologically active portion of an MP protein" is intended to include a portion, *e.g.*, a domain/motif, of an MP protein that catalyzes an enzymatic reaction in one or more *C. glutamicum* amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathways, or has an activity as set forth in Table 1. To determine whether an MP protein or a biologically active portion thereof can catalyze an enzymatic reaction in an amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathway, an assay of enzymatic activity may be performed. Such assay methods are well known to those of ordinary skill in the art, as detailed in Example 8 of the Exemplification.

Additional nucleic acid fragments encoding biologically active portions of an MP protein can be prepared by isolating a portion of one of the amino acid sequences of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing), expressing the encoded portion of the MP protein or peptide (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the MP protein or peptide.

The invention further encompasses nucleic acid molecules that differ from one of the nucleotide sequences of the invention (*e.g.*, a sequence of an odd-numbered SEQ ID NO: of the Sequence Listing) (and portions thereof) due to degeneracy of the genetic code and thus encode the same MP protein as that encoded by the nucleotide sequences of the invention. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in the Sequence Listing (*e.g.*, an even-numbered SEQ ID NO:). In a still further embodiment, the nucleic acid molecule of the invention encodes a full length *C. glutamicum* protein which is substantially homologous to an amino acid sequence of the invention (encoded by an open reading frame shown in an odd-numbered SEQ ID NO: of the Sequence Listing).

It will be understood by one of ordinary skill in the art that in one embodiment the sequences of the invention are not meant to include the sequences of the prior art, such as those Genbank sequences set forth in Table 3 which were available prior to the present invention. In one embodiment, the invention includes

nucleotide and amino acid sequences having a percent identity to a nucleotide or amino acid sequence of the invention which is greater than that of a sequence of the prior art, i.e the invention includes a nucleotide sequence which encodes a proteine sequence which is greater than and/or at least 71% identical to the proteine sequence designated SEQ ID NO:2 and/or a nucleotide sequence which encodes a proteine sequence which is greater than and/or at least 63 % identical to the proteine sequence designated SEQ ID NO: 4.

Table 3: Alignment results

Gene name (identifier)	Genbank hit	Homology	Reference
metH	GB_BA2:MTCY261 Mycobacterium tuberculosis H37Rv Complete genome	70.3 %	Cole et al. (1998) Nature 393, 537–544
treS	GB_BA2:MTCY261 Mycobacterium tuberculosis H37Rv complete genome	62.4 %	Cole et al. (1998) Nature 393, 537–544

Homology: CLUSTAL-calculated percent identity (Open reading frames from the genome, translated into amino acid sequence)

One of ordinary skill in the art would be able to calculate the lower threshold of percent identity for any given sequence of the invention by examining the CLUSTAL-calculated percent identity scores set forth in Table 3 for each of the three top hits for the given sequence. One of ordinary skill in the art will also appreciate that nucleic acid and amino acid sequences having percent identities greater than the lower threshold so calculated (*e.g.*, preferably at least about 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more identical) are also encompassed by the invention.

In addition to the *C. glutamicum* MP nucleotide sequences set forth in the Sequence Listing as odd-numbered SEQ ID NOs, it will be appreciated by one of ordinary skill in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of MP proteins may exist within a population (*e.g.*, the

C. glutamicum population). Such genetic polymorphism in the MP gene may exist among individuals within a population due to natural variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding an MP protein, preferably a *C. glutamicum* MP protein. Such natural variations can typically result in 1-5% variance in the nucleotide sequence of the MP gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in MP that are the result of natural variation and that do not alter the functional activity of MP proteins are intended to be within the scope of the invention.

Nucleic acid molecules corresponding to natural variants and non-*C. glutamicum* homologues of the *C. glutamicum* MP DNA of the invention can be isolated based on their homology to the *C. glutamicum* MP nucleic acid disclosed herein using the *C. glutamicum* DNA, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising a nucleotide sequence of an odd-numbered SEQ ID NO: of the Sequence Listing. In other embodiments, the nucleic acid is at least 30, 50, 100, 250 or more nucleotides in length. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are such that sequences at least about 65%, more preferably at least about 70%, and even more preferably at least about 75% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to one of ordinary skill in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to a nucleotide sequence of the invention corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein). In one embodiment, the nucleic acid encodes a natural *C. glutamicum* MP protein.

In addition to naturally-occurring variants of the MP sequence that may exist in the population, one of ordinary skill in the art will further appreciate that changes can be introduced by mutation into a nucleotide sequence of the invention, thereby leading to changes in the amino acid sequence of the encoded MP protein, without altering the functional ability of the MP protein. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in a nucleotide sequence of the invention. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of one of the MP proteins (*e.g.*, an even-numbered SEQ ID NO: of the Sequence Listing) without altering the activity of said MP protein, whereas an "essential" amino acid residue is required for MP protein activity. Other amino acid residues, however, (*e.g.*, those that are not conserved or only semi-conserved in the domain having MP activity) may not be essential for activity and thus are likely to be amenable to alteration without altering MP activity.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding MP proteins that contain changes in amino acid residues that are not essential for MP activity. Such MP proteins differ in amino acid sequence from a sequence of an even-numbered SEQ ID NO: of the Sequence Listing yet retain at least one of the MP activities described herein. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 50% homologous to an amino acid sequence of the invention and is capable of catalyzing an enzymatic reaction in an amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathway, or has one or more activities set forth in Table 1. Preferably, the protein encoded by the nucleic acid molecule is at least about 50-60% homologous to the amino acid sequence of one of the odd-numbered SEQ ID NOs of the Sequence Listing, more preferably at least about 60-70% homologous to one of these sequences, even more preferably at least about 70-80%, 80-90%, 90-95% homologous to one of these sequences, and most preferably at least about 96%, 97%, 98%, or 99% homologous to one of the amino acid sequences of the invention.

To determine the percent homology of two amino acid sequences (*e.g.*, one of the amino acid sequences of the invention and a mutant form thereof) or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of one protein or nucleic acid for optimal alignment with the other protein or nucleic acid). The

amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in one sequence (*e.g.*, one of the amino acid sequences of the invention) is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence (*e.g.*, a mutant form of the amino acid sequence), then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity"). The percent homology between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % homology = # of identical positions/total # of positions x 100).

An isolated nucleic acid molecule encoding an MP protein homologous to a protein sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing) can be created by introducing one or more nucleotide substitutions, additions or deletions into a nucleotide sequence of the invention such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into one of the nucleotide sequences of the invention by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in an MP protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an MP coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for an MP activity described herein to identify mutants that retain MP activity. Following mutagenesis of the nucleotide sequence of one of the odd-numbered SEQ ID NOs of the Sequence Listing, the encoded protein can be expressed recombinantly and the activity of the protein can be

determined using, for example, assays described herein (see Example 8 of the Exemplification).

In addition to the nucleic acid molecules encoding MP proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded DNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire MP coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an MP protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding MP. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding MP disclosed herein (*e.g.*, the sequences set forth as odd-numbered SEQ ID NOs in the Sequence Listing), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of MP mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of MP mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of MP mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil,

5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5- oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a cell or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an MP protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. The antisense molecule can be modified such that it specifically binds to a receptor or an antigen expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecule to a peptide or an antibody which binds to a cell surface receptor or antigen. The antisense nucleic acid molecule can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong prokaryotic, viral, or eukaryotic promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An -anomeric nucleic acid molecule forms specific double-stranded

hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave MP mRNA transcripts to thereby inhibit translation of MP mRNA. A ribozyme having specificity for an MP-encoding nucleic acid can be designed based upon the nucleotide sequence of an MP DNA disclosed herein (*i.e.*, SEQ ID NO: 1 (RXA02229). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an MP-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071 and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, MP mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, MP gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of an MP nucleotide sequence (*e.g.*, an MP promoter and/or enhancers) to form triple helical structures that prevent transcription of an MP gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

Another aspect of the invention pertains combinations of gene in the methionine and/or lysine metabolism. Preferred combinations are the combination of metZ with metC, metB (encoding Cystathionine-Synthase), metA (encoding homoserine-O-acetyltransferase), metE (encoding Methionine Synthase), metH (encoding Methionine Synthase, herein designated as SEQ ID No: 1), hom (encoding homoserine dehydrogenase), asd (encoding aspartatesemialdehyd dehydrogenase), ask (encoding aspartokinase) and rxa00657 (table 4).

Table 4

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Gene name (identifier)	Function
5	6	MetZ	Acetylhomoserine sulfhydrolase
7	8	RXA00657	

It may be that all of the genes are expressed in a host strain. But it is also possible that only a part of the mentioned genes is chosen, e.g. metZ and metA, or metZ, metA, metH and hom or any other of the possible combinations. The genes may be altered in their nucleotide and in the corresponding amino acid sequence resulting in derivatives in such a way that their activity is altered under physiological conditions which leads to an increase in productivity and/or yield of a desired fine chemical. One class of such alterations or derivatives is well known for the nucleotide sequence of the ask gene encoding aspartokinase. These alterations lead to removal of feed back inhibition by the amino acids lysine and threonine and subsequently to lysine overproduction. In a preferred embodiment the metH gene or altered forms of the metH gene are used in a Corynebacterium strain in combination with ask, hom, metA and metZ or derivatives of these genes. In another preferred embodiment metH or altered forms of the metH gene are used in a Corynebacterium strain in combination with ask, hom, metA, metZ and metE or derivatives of these genes. In a more preferred embodiment the gene combinations metH or altered forms of the metH gene are combined with ask, hom, metA and metZ or derivatives of these genes, or metH is combined with ask, hom, metA, metZ and metE or derivatives of these genes in a Corynebacterium strain and sulfur sources like sulfates, thiosulfates, sulfites and also more reduced sulfur sources like H_2S and sulfides and derivatives are used in the growth medium. Also sulfur sources like methyl mercaptan, methanesulfonic acid, thioglycolates, thiocyanates, thiourea, sulfur containing amino acids like cysteine and other sulfur containing compounds can be fed. Another aspect of the invention pertains to the use of the above mentioned gene combinations in a Corynebacterium strain wich is before or after introduction of the genes mutagenized by radiation or by well known mutagenic chemicals and selected for resistancy against high concentrations of the fine chmical of interest, e.g. lysine or methionine or anologes of the desired fine chemical like the methionine analogons ethionine or methyl methionine or others. In another embodiment the gene combinations mentioned above can be expressed in a Corynebacterium strain having particular gene disruptions. Preferred are gene diruptions that encode proteins that favor

carbon flux to undesired metabolites. In case methionine is the desired fine chemical the formation of lysine may be unfavorable. In such a case the combination of the above mentioned genes should proceed in a Corynebacterium strain bearing a gene disruption of the lysA gene (encoding diaminopimelate decarboxylase) or the ddh gene (encoding the meso-diaminopimelate dehydrogenase catalysing the conversion of tetrahydropicolinate to meso-diaominopimelate). In a preferred embodiment a favorable combination of the above mentioned genes are all altered in such a way that their gene products are not feed back inhibited by endproducts or metabolites of the biosynthetic pathway leading to the desired fine chemical. In the case that the desired fine chemical is methionine, the gene combinations may be expressed in a strain previously treated with mutagenic agents or radiation and selected for the above mentioned resistancies. Additionally the strain should be grown in a growth medium containing one or more of the above mentioned sulfur sources.

Another aspect of the invention pertains combinations of genes involved in the metabolism of trehalose and the combination of genes involved in the metabolism of trehalose and other mono-, oligo- or polymeric saccharides. Preferred are combinations of the gene for trehalose synthase (herein designated as SEQ ID No: 3) with genes disclosed in table 5.

Another aspect of the invention is the combination of the gene for trehalose synthase with genes involved in saccharide import, as e.g. the genes for the PTS system (as disclosed in table 5), other saccharide transport systems or proteins facilitating saccharide efflux from the cell into the surrounding environment.

TABLE 5: PHOSPHOENOLPYRUVATE: SUGAR PHOSPHOTRANSFERASE SYSTEM

Nucleotide SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Function
9	10	RXS00315	PTS SYSTEM, SUCROSE–SPECIFIC IIABC COMPONENT (EIIABC–SCR) (SUCROSE– PERMEASE IIABC COMPONENT(PHOSPHO-TRANSFERASE ENZYME II, ABC COMPONENT) (EC 2.7.1.69)
11	12	RXN01299	PTS SYSTEM, FRUCTOSE–SPECIFIC IIBC COMPONENT (EC 2.7.1.69)
13	14	RXA00951	PTS SYSTEM, MANNITOL (CRYPTIC) –SPECIFIC IIA COMPONENT (EIIA–(C)MTL) (MANNITOL (CRYPTIC)– PERMEASE IIA COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT) (EC 2.7.1.69)
15	16	RXN01244	PHOSPHOENOLPYRUVATE–PROTEIN PHOSPHOTRANSFERASE (EC 2.7.3.9)
17	18	RXA01300	PHOSPHOCARRIER PROTEIN HPR
19	20	RXN03002	PTS SYSTEM, MANNITOL (CRYPTIC) –SPECIFIC IIA COMPONENT (EIIA–(C)MTL) (MANNITOL (CRYPTIC)–PERMEASE IIA COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT) (EC 2.7.1.69)
21	22	RXC00953	Membrane Spanning Protein involved in PTS system
23	24	RXC03001	Membrane Spanning Protein involved in PTS system
25	26	RXN01943	PTS SYSTEM, GLUCOSE–SPECIFIC IIABC COMPONENT (EC 2.7.1.69)
27	28	RXA01503	PTS SYSTEM, BETA–GLUCOSIDES–SPECIFIC IIABC COMPONENT (EIIABC–BGL) (BETA–GLUCOSIDES– PERMEASE IIABC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, ABC COMPONENT) (EC 2.7.1.69)

Trehalose

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Function
29	30	RXN00351	ALPHA,ALPHA–TREHALOSE–PHOSPHATE SYNTHASE (UDP–FORMING) 56 KD SUBUNIT (EC 2.4.1.15)
31	32	RXA00347	TREHALOSE–PHOSPHATASE (EC 3.1.3.12)
33	34	RXN01239	maltooligosyltrehalose synthase
35	36	RXA02645	maltooligosyltrehalose trehalohydrolase
37	38	RXN02355	TREHALOSE/MALTOSE BINDING PROTEIN
39	40	RXN02909	Hypothetical Trehalose–Binding Protein
41	42	RXS00349	Hypothetical Trehalose Transport Protein
43	44	RXS03183	TREHALOSE/MALTOSE BINDING PROTEIN
45	46	RXC00874	transmebrane protein involved in trehalose metabolism

Another aspect of the invention pertains to the use of the above mentioned gene combinations in a Corynebacterium strain wich is before or after introduction of the genes mutagenized by radiation or by well known mutagenic chemicals and selected for

resistancy against high concentrations of feedstock (as e.g. glucose or other saccharides) or the fine chemical of interest, e.g. trehalose or other saccharides.

In another embodiment the gene combinations mentioned above can be expressed in a Corynebacterium strain having particular gene disruptions or gene attenuations (i.e. genes which biological activity is reduced compared to the normal level). Preferred are disruptions or attenuations of genes that encode proteins that favor carbon flux to metabolic pathways which do not lead to the desired fine chemical. In case of trehalose being the desired fine chemical, such less desired metabolic pathways may be e.g. glycolysis or pentose phosphate cycle.

B. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an MP protein (or a portion thereof) or combinations of genes wherein at least one gene encodes for an MP protein. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the

recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, repressor binding sites, activator binding sites, enhancers and other expression control elements (*e.g.*, terminators, polyadenylation signals, or other elements of mRNA secondary structure). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells. Preferred regulatory sequences are, for example, promoters such as cos-, tac-, trp-, tet-, trp-tet-, lpp-, lac-, lpp-lac-, lacI^q-, T7-, T5-, T3-, gal-, trc-, ara-, SP6-, arny, SPO2, ë-P_R- or ë P_L, which are used preferably in bacteria. Additional regulatory sequences are, for example, promoters from yeasts and fungi, such as ADC1, MFá, AC, P-60, CYC1, GAPDH, TEF, rp28, ADH, promoters from plants such as CaMV/35S, SSU, OCS, lib4, usp, STLS1, B33, nos or ubiquitin- or phaseolin-promoters. It is also possible to use artificial promoters. It will be appreciated by one of ordinary skill in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, MP proteins, mutant forms of MP proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of MP proteins in prokaryotic or eukaryotic cells. For example, MP genes can be expressed in bacterial cells such as *C. glutamicum*, insect cells (using baculovirus expression vectors), yeast and other fungal cells (see Romanos, M.A. *et al.* (1992) "Foreign gene expression in yeast: a review", *Yeast* 8: 423-488; van den Hondel, C.A.M.J.J. *et al.* (1991) "Heterologous gene expression in filamentous fungi" in: More Gene Manipulations in Fungi, J.W. Bennet & L.L. Lasure, eds., p. 396-428: Academic Press: San Diego; and van den Hondel,

C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, Peberdy, J.F. *et al.*, eds., p. 1-28, Cambridge University Press: Cambridge), algae and multicellular plant cells (see Schmidt, R. and Willmitzer, L. (1988) High efficiency *Agrobacterium tumefaciens* -mediated transformation of *Arabidopsis thaliana* leaf and cotyledon explants" *Plant Cell Rep.*: 583-586), or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein but also to the C-terminus or fused within suitable regions in the proteins. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. In one embodiment, the coding sequence of the MP protein is cloned into a pGEX expression vector to create a vector encoding a fusion protein comprising, from the N-terminus to the C-terminus, GST-thrombin cleavage site-X protein. The fusion protein can be purified by affinity chromatography using glutathione-agarose resin. Recombinant MP protein unfused to GST can be recovered by cleavage of the fusion protein with thrombin.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) pLG338, pACYC184, pBR322, pUC18, pUC19, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III113-B1, ëgt11, pBdCl, and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89; and Pouwels *et al.*, eds. (1985) Cloning Vectors. Elsevier: New York IBSN 0 444 904018). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter. For transformation of other varieties of bacteria, appropriate vectors may be selected. For example, the plasmids pIJ101, pIJ364, pIJ702 and pIJ361 are known to be useful in transforming Streptomyces, while plasmids pUB110, pC194, or pBD214 are suited for transformation of Bacillus species. Several plasmids of use in the transfer of genetic information into Corynebacterium include pHM1519, pBL1, pSA77, or pAJ667 (Pouwels *et al.*, eds. (1985) Cloning Vectors. Elsevier: New York IBSN 0 444 904018).

One strategy to maximize recombinant protein expression is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in the bacterium chosen for expression, such as *C. glutamicum* (Wada *et al.* (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the MP protein expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), , 2 ì, pAG-1, Yep6, Yep13, pEMBLYe23, pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Vectors and methods for the construction of vectors appropriate for use in other fungi, such as the filamentous fungi, include those detailed in: van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and

vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, J.F. Peberdy, *et al.*, eds., p. 1-28, Cambridge University Press: Cambridge, and Pouwels *et al.*, eds. (1985) Cloning Vectors. Elsevier: New York (IBSN 0 444 904018).

Alternatively, the MP proteins of the invention can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In another embodiment, the MP proteins of the invention may be expressed in unicellular plant cells (such as algae) or in plant cells from higher plants (*e.g.*, the spermatophytes, such as crop plants). Examples of plant expression vectors include those detailed in: Becker, D., Kemper, E., Schell, J. and Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20: 1195-1197; and Bevan, M.W. (1984) "Binary *Agrobacterium* vectors for plant transformation", *Nucl. Acid. Res.* 12: 8711-8721, and include pLGV23, pGHlac+, pBIN19, pAK2004, and pDH51 (Pouwels *et al.*, eds. (1985) Cloning Vectors. Elsevier: New York IBSN 0 444 904018).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular

promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to MP mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, an MP protein can be expressed in bacterial cells such as *C. glutamicum*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those of ordinary skill in the art. Microorganisms related to *Corynebacterium glutamicum* which may be conveniently used as host cells for the nucleic acid and protein molecules of the invention are set forth in Table 2.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection", "conjugation" and "transduction" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, linear DNA or RNA (*e.g.*, a linearized vector or a gene construct alone without a vector) or nucleic acid in the form of a vector (*e.g.*, a plasmid, phage, phasmid, phagemid, transposon or other DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, natural competence, chemical-mediated transfer, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding an MP protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

To create a homologous recombinant microorganism, a vector is prepared which contains at least a portion of an MP gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the MP gene.

Preferably, this MP gene is a *Corynebacterium glutamicum* MP gene, but it can be a homologue from a related bacterium or even from a mammalian, yeast, or insect source. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous MP gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous MP gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous MP protein). In the homologous recombination vector, the altered portion of the MP gene is flanked at its 5' and 3' ends by additional nucleic acid of the MP gene to allow for homologous recombination to occur between the exogenous MP gene carried by the vector and an endogenous MP gene in a microorganism. The additional flanking MP nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see *e.g.*, Thomas, K.R., and Capecchi, M.R. (1987) Cell 51: 503 for a description of homologous recombination vectors). The vector is introduced into a microorganism (*e.g.*, by electroporation) and cells in which the introduced MP gene has homologously recombined with the endogenous MP gene are selected, using art-known techniques.

In another embodiment, recombinant microorganisms can be produced which contain selected systems which allow for regulated expression of the introduced gene. For example, inclusion of an MP gene on a vector placing it under control of the lac operon permits expression of the MP gene only in the presence of IPTG. Such regulatory systems are well known in the art.

In another embodiment, an endogenous MP gene in a host cell is disrupted (*e.g.*, by homologous recombination or other genetic means known in the art) such that expression of its protein product does not occur. In another embodiment, an endogenous or introduced MP gene in a host cell has been altered by one or more point mutations, deletions, or inversions, but still encodes a functional MP protein. In still another embodiment, one or more of the regulatory regions (*e.g.*, a promoter, repressor, or inducer) of an MP gene in a microorganism has been altered (*e.g.*, by deletion, truncation, inversion, or point mutation) such that the expression of the MP gene is modulated. One of ordinary skill in the art will appreciate that host cells containing more than one of the described MP gene and protein modifications may be

readily produced using the methods of the invention, and are meant to be included in the present invention.

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) an MP protein. Accordingly, the invention further provides methods for producing MP proteins using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding an MP protein has been introduced, or into which genome has been introduced a gene encoding a wild-type or altered MP protein) in a suitable medium until MP protein is produced. In another embodiment, the method further comprises isolating MP proteins from the medium or the host cell.

C. Isolated MP Proteins

Another aspect of the invention pertains to isolated MP proteins, and biologically active portions thereof. An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of MP protein in which the protein is separated from cellular components of the cells in which it is naturally or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of MP protein having less than about 30% (by dry weight) of non-MP protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-MP protein, still more preferably less than about 10% of non-MP protein, and most preferably less than about 5% non-MP protein. When the MP protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation. The language "substantially free of chemical precursors or other chemicals" includes preparations of MP protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of MP protein having less than about 30% (by dry weight) of chemical precursors or non-MP chemicals, more preferably less than about 20% chemical precursors or non-MP chemicals, still more preferably less than about 10% chemical

precursors or non-MP chemicals, and most preferably less than about 5% chemical precursors or non-MP chemicals. In preferred embodiments, isolated proteins or biologically active portions thereof lack contaminating proteins from the same organism from which the MP protein is derived. Typically, such proteins are produced by recombinant expression of, for example, a *C. glutamicum* MP protein in a microorganism such as *C. glutamicum*.

An isolated MP protein or a portion thereof of the invention can catalyze an enzymatic reaction in an amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathway, or has one or more of the activities set forth in Table 1. In preferred embodiments, the protein or portion thereof comprises an amino acid sequence which is sufficiently homologous to an amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing) such that the protein or portion thereof maintains the ability to catalyze an enzymatic reaction in an amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathway. The portion of the protein is preferably a biologically active portion as described herein. In another preferred embodiment, an MP protein of the invention has an amino acid sequence set forth as an even-numbered SEQ ID NO: of the Sequence Listing. In yet another preferred embodiment, the MP protein has an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, to a nucleotide sequence of the invention (*e.g.*, a sequence of an odd-numbered SEQ ID NO: of the Sequence Listing). In still another preferred embodiment, the MP protein has an amino acid sequence which is encoded by a nucleotide sequence that is preferably at least about 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to one of the nucleic acid sequences of the invention, or a portion thereof. Ranges and identity values intermediate to the above-recited values, (*e.g.*, 70-90% identical or 80-95% identical) are also intended to be encompassed by the present invention. For example, ranges of identity values using a combination of any of the above values recited as upper and/or lower limits are intended to be included. The preferred MP proteins of the present invention also preferably possess at least one of the MP activities described herein. For example, a preferred MP protein of the present invention includes an amino acid sequence encoded by a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, to a nucleotide

sequence of the invention, and which can catalyze an enzymatic reaction in an amino acid, vitamin, cofactor, nutraceutical, nucleotide, nucleoside, or trehalose metabolic pathway, or which has one or more of the activities set forth in Table 1.

In other embodiments, the MP protein is substantially homologous to an amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing) and retains the functional activity of the protein of one of the amino acid sequences of the invention yet differs in amino acid sequence due to natural variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the MP protein is a protein which comprises an amino acid sequence which is preferably at least about 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to an entire amino acid sequence of the invention and which has at least one of the MP activities described herein. Ranges and identity values intermediate to the above-recited values, (*e.g.*, 70-90% identical or 80-95% identical) are also intended to be encompassed by the present invention. For example, ranges of identity values using a combination of any of the above values recited as upper and/or lower limits are intended to be included. In another embodiment, the invention pertains to a full length *C. glutamicum* protein which is substantially homologous to an entire amino acid sequence of the invention.

Biologically active portions of an MP protein include peptides comprising amino acid sequences derived from the amino acid sequence of an MP protein, *e.g.*, an amino acid sequence of an even-numbered SEQ ID NO: of the Sequence Listing or the amino acid sequence of a protein homologous to an MP protein, which include fewer amino acids than a full length MP protein or the full length protein which is homologous to an MP protein, and exhibit at least one activity of an MP protein. Typically, biologically active portions (peptides, *e.g.*, peptides which are, for example, 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) comprise a domain or motif with at least one activity of an MP protein. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the activities described herein. Preferably, the biologically active portions of an MP protein include one

or more selected domains/motifs or portions thereof having biological activity.

MP proteins are preferably produced by recombinant DNA techniques. For example, a nucleic acid molecule encoding the protein is cloned into an expression vector (as described above), the expression vector is introduced into a host cell (as described above) and the MP protein is expressed in the host cell. The MP protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Alternative to recombinant expression, an MP protein, polypeptide, or peptide can be synthesized chemically using standard peptide synthesis techniques. Moreover, native MP protein can be isolated from cells (*e.g.*, endothelial cells), for example using an anti-MP antibody, which can be produced by standard techniques utilizing an MP protein or fragment thereof of this invention.

The invention also provides MP chimeric or fusion proteins. As used herein, an MP "chimeric protein" or "fusion protein" comprises an MP polypeptide operatively linked to a non-MP polypeptide. An "MP polypeptide" refers to a polypeptide having an amino acid sequence corresponding to MP, whereas a "non-MP polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the MP protein, *e.g.*, a protein which is different from the MP protein and which is derived from the same or a different organism. Within the fusion protein, the term "operatively linked" is intended to indicate that the MP polypeptide and the non-MP polypeptide are fused in-frame to each other. The non-MP polypeptide can be fused to the N-terminus or C-terminus of the MP polypeptide. For example, in one embodiment the fusion protein is a GST-MP fusion protein in which the MP sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant MP proteins. In another embodiment, the fusion protein is an MP protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of an MP protein can be increased through use of a heterologous signal sequence.

Preferably, an MP chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for

appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). An MP-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the MP protein.

Homologues of the MP protein can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of the MP protein. As used herein, the term "homologue" refers to a variant form of the MP protein which acts as an agonist or antagonist of the activity of the MP protein. An agonist of the MP protein can retain substantially the same, or a subset, of the biological activities of the MP protein. An antagonist of the MP protein can inhibit one or more of the activities of the naturally occurring form of the MP protein, by, for example, competitively binding to a downstream or upstream member of the MP cascade which includes the MP protein. Thus, the *C. glutamicum* MP protein and homologues thereof of the present invention may modulate the activity of one or more metabolic pathways in which MP proteins play a role in this microorganism.

In an alternative embodiment, homologues of the MP protein can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the MP protein for MP protein agonist or antagonist activity. In one embodiment, a variegated library of MP variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of MP variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential MP sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of MP sequences therein. There are a variety of methods which can be used to produce libraries of potential MP homologues from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic

DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential MP sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477.

In addition, libraries of fragments of the MP protein coding can be used to generate a variegated population of MP fragments for screening and subsequent selection of homologues of an MP protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an MP coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the MP protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of MP homologues. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify MP homologues (Arkin and Yourvan (1992) *PNAS 89*:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

In another embodiment, cell based assays can be exploited to analyze a variegated MP library, using methods well known in the art.

D. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, fusion proteins, primers, vectors, and host cells described herein can be used in one or more of the following methods: identification of *C. glutamicum* and related organisms; mapping of genomes of organisms related to *C. glutamicum*; identification and localization of *C. glutamicum* sequences of interest; evolutionary studies; determination of MP protein regions required for function; modulation of an MP protein activity; modulation of the activity of an MP pathway; and modulation of cellular production of a desired compound, such as a fine chemical. The MP nucleic acid molecules of the invention have a variety of uses. First, they may be used to identify an organism as being *Corynebacterium glutamicum* or a close relative thereof. Also, they may be used to identify the presence of *C. glutamicum* or a relative thereof in a mixed population of microorganisms. The invention provides the nucleic acid sequences of a number of *C. glutamicum* genes; by probing the extracted genomic DNA of a culture of a unique or mixed population of microorganisms under stringent conditions with a probe spanning a region of a *C. glutamicum* gene which is unique to this organism, one can ascertain whether this organism is present. Although *Corynebacterium glutamicum* itself is not pathogenic to humans, it is related to species which are human pathogens, such as *Corynebacterium diphtheriae*. *Corynebacterium diphtheriae* is the causative agent of diphtheria, a rapidly developing, acute, febrile infection which involves both local and systemic pathology. In this disease, a local lesion develops in the upper respiratory tract and involves necrotic injury to epithelial cells; the bacilli secrete toxin which is disseminated through this lesion to distal susceptible tissues of the body. Degenerative changes brought about by the inhibition of protein synthesis in these tissues, which include heart, muscle, peripheral nerves, adrenals, kidneys, liver and spleen, result in the systemic pathology of the disease. Diphtheria continues to have high incidence in many parts of the world, including Africa, Asia, Eastern Europe and the independent states of the former Soviet Union. An ongoing epidemic of diphtheria in the latter two regions has resulted in at least 5,000 deaths since 1990.

In one embodiment, the invention provides a method of identifying the presence or activity of *Cornyebacterium diphtheriae* in a subject. This method includes detection of one or more of the

nucleic acid or amino acid sequences of the invention (*e.g.*, the sequences set forth as odd-numbered or even-numbered SEQ ID NOs, respectively, in the Sequence Listing) in a subject, thereby detecting the presence or activity of *Corynebacterium diphtheriae* in the subject. *C. glutamicum* and *C. diphtheriae* are related bacteria, and many of the nucleic acid and protein molecules in *C. glutamicum* are homologous to *C. diphtheriae* nucleic acid and protein molecules, and can therefore be used to detect *C. diphtheriae* in a subject.

The nucleic acid and protein molecules of the invention may also serve as markers for specific regions of the genome. This has utility not only in the mapping of the genome, but also for functional studies of *C. glutamicum* proteins. For example, to identify the region of the genome to which a particular *C. glutamicum* DNA-binding protein binds, the *C. glutamicum* genome could be digested, and the fragments incubated with the DNA-binding protein. Those which bind the protein may be additionally probed with the nucleic acid molecules of the invention, preferably with readily detectable labels; binding of such a nucleic acid molecule to the genome fragment enables the localization of the fragment to the genome map of *C. glutamicum*, and, when performed multiple times with different enzymes, facilitates a rapid determination of the nucleic acid sequence to which the protein binds. Further, the nucleic acid molecules of the invention may be sufficiently homologous to the sequences of related species such that these nucleic acid molecules may serve as markers for the construction of a genomic map in related bacteria, such as *Brevibacterium lactofermentum*.

The MP nucleic acid molecules of the invention are also useful for evolutionary and protein structural studies. The metabolic processes in which the molecules of the invention participate are utilized by a wide variety of prokaryotic and eukaryotic cells; by comparing the sequences of the nucleic acid molecules of the present invention to those encoding similar enzymes from other organisms, the evolutionary relatedness of the organisms can be assessed. Similarly, such a comparison permits an assessment of which regions of the sequence are conserved and which are not, which may aid in determining those regions of the protein which are essential for the functioning of the enzyme. This type of determination is of value for protein engineering studies and may give an indication of what the protein can tolerate in terms of mutagenesis without losing function.

Manipulation of the MP nucleic acid molecules of the invention may result in the production of MP proteins having functional differences from the wild-type MP proteins. These proteins may be improved in efficiency or activity, may be present in greater numbers in the cell than is usual, or may be decreased in efficiency or activity.

The invention also provides methods for screening molecules which modulate the activity of an MP protein, either by interacting with the protein itself or a substrate or binding partner of the MP protein, or by modulating the transcription or translation of an MP nucleic acid molecule of the invention. In such methods, a microorganism expressing one or more MP proteins of the invention is contacted with one or more test compounds, and the effect of each test compound on the activity or level of expression of the MP protein is assessed.

When the desired fine chemical to be isolated from large-scale fermentative culture of *C. glutamicum* is an amino acid, a vitamin, a cofactor, a nutraceutical, a nucleotide, a nucleoside, or trehalose, modulation of the activity or efficiency of activity of one or more of the proteins of the invention by recombinant genetic mechanisms may directly impact the production of one of these fine chemicals. For example, in the case of an enzyme in a biosynthetic pathway for a desired amino acid, improvement in efficiency or activity of the enzyme (including the presence of multiple copies of the gene) should lead to an increased production or efficiency of production of that desired amino acid. In the case of an enzyme in a biosynthetic pathway for an amino acid whose synthesis is in competition with the synthesis of a desired amino acid, any decrease in the efficiency or activity of this enzyme (including deletion of the gene) should result in an increase in production or efficiency of production of the desired amino acid, due to decreased competition for intermediate compounds and/or energy. In the case of an enzyme in a degradation pathway for a desired amino acid, any decrease in efficiency or activity of the enzyme should result in a greater yield or efficiency of production of the desired product due to a decrease in its degradation. Lastly, mutagenesis of an enzyme involved in the biosynthesis of a desired amino acid such that this enzyme is no longer is capable of feedback inhibition should result in increased yields or efficiency of production of the desired amino acid. The same should apply to the biosynthetic and degradative enzymes of the invention involved in the metabolism of vitamins, cofactors, nutraceuticals, nucleotides, nucleosides and trehalose.

Similarly, when the desired fine chemical is not one of the aforementioned compounds, the modulation of activity of one of the proteins of the invention may still impact the yield and/or efficiency of production of the compound from large-scale culture of *C. glutamicum*. The metabolic pathways of any organism are closely interconnected; the intermediate used by one pathway is often supplied by a different pathway. Enzyme expression and function may be regulated based on the cellular levels of a compound from a different metabolic process, and the cellular levels of molecules necessary for basic growth, such as amino acids and nucleotides, may critically affect the viability of the microorganism in large-scale culture. Thus, modulation of an amino acid biosynthesis enzyme, for example, such that it is no longer responsive to feedback inhibition or such that it is improved in efficiency or turnover may result in increased cellular levels of one or more amino acids. In turn, this increased pool of amino acids provides not only an increased supply of molecules necessary for protein synthesis, but also of molecules which are utilized as intermediates and precursors in a number of other biosynthetic pathways. If a particular amino acid had been limiting in the cell, its increased production might increase the ability of the cell to perform numerous other metabolic reactions, as well as enabling the cell to more efficiently produce proteins of all kinds, possibly increasing the overall growth rate or survival ability of the cell in large scale culture. Increased viability improves the number of cells capable of producing the desired fine chemical in fermentative culture, thereby increasing the yield of this compound. Similar processes are possible by the modulation of activity of a degradative enzyme of the invention such that the enzyme no longer catalyzes, or catalyzes less efficiently, the degradation of a cellular compound which is important for the biosynthesis of a desired compound, or which will enable the cell to grow and reproduce more efficiently in large-scale culture. It should be emphasized that optimizing the degradative activity or decreasing the biosynthetic activity of certain molecules of the invention may also have a beneficial effect on the production of certain fine chemicals from *C. glutamicum*. For example, by decreasing the efficiency of activity of a biosynthetic enzyme in a pathway which competes with the biosynthetic pathway of a desired compound for one or more intermediates, more of those intermediates should be available for conversion to the desired product. A similar situation may call for the improvement of degradative ability or efficiency of one or more proteins of the invention.

This aforementioned list of mutagenesis strategies for MP proteins to result in increased yields of a desired compound is not meant to be limiting; variations on these mutagenesis strategies will be readily apparent to one of ordinary skill in the art. By these mechanisms, the nucleic acid and protein molecules of the invention may be utilized to generate *C. glutamicum* or related strains of bacteria expressing mutated MP nucleic acid and protein molecules such that the yield, production, and/or efficiency of production of a desired compound is improved. This desired compound may be any natural product of *C. glutamicum*, which includes the final products of biosynthesis pathways and intermediates of naturally-occurring metabolic pathways, as well as molecules which do not naturally occur in the metabolism of *C. glutamicum*, but which are produced by a *C. glutamicum* strain of the invention. Preferred compounds to be produced by Corynebacterium glutamicum strains are trehalose and/or the amino acids L-lysine and L-methionine.

In one embodiment the *metC* gene encoding cystathionine â-lyase, the third enzyme in the methionine biosynthetic pathway, was isolated from *Corynebacterium glutamicum*. The translational product of the gene showed no significant homology with that of *metC* gene from other organisms. Introduction of the plasmid containing the *metC* gene into *C. glutamicum* resulted in 5-fold increase in the activity of cystathionine â-lyase. The protein product now designated MetC encoding a protein product of 35,574 Dalton consisted of 325 amino acids was identical to the previously reported *aecD* gene except the existence of two different amino acids. Like *aecD* gene, when present in multiple copies, *metC* gene conferred resistance to *S*-(â-aminoethyl)-cysteine which is a toxic lysine analog. However, genetic and biochemical evidences suggest that the natural activity of *metC* gene product is to mediate methionine biosynthesis in *C. glutamicum*. Mutant strains of *metC* were constructed and the strains showed methionine prototrophy. The mutant strains completely lost their ability to show resistance to *S*-(ã-aminoethyl)-cysteine. These results show that, in addition to the transsulfuration, another biosynthetic pathway - the direct sulfhydrylation pathway is functional in *C. glutamicum* as a parallel biosynthetic route for methionine.

In yet another embodiment it is also shown that the additional sulfhydrylation pathway is catalyzed by *O*-acetylhomoserine sulfhydrylase. The presence of the pathway is demonstrated by the isolation of the corresponding *metZ* (or *metY*) gene and enzyme. Among the eukaryotes, fungi and yeast species have been reported to have both the transsulfuration and direct sulfhydrylation

pathway (Marzluf, 1997). So far, no prokaryotic organism which possesses both pathways has been found. Unlike *E. coli* which only possesses single biosynthetic route for lysine, *C. glutamicum* possesses two parallel biosynthetic pathways for the amino acid. The biosynthetic pathway for methionine in *C. glutamicum* is analogous to that of lysine in that aspect.

The Gene *metZ* was found because it was located in the upstream region of *metA*. We sequenced regions upstream and downstream of *metA* - the gene encoding the enzyme catalysing the first step of methionine biosynthesis (Park, S.-D., Lee, J.-Y., Kim, Y., Kim, J.-H., and Lee, H.-S. (1998) Isolation and analysis of *metA*, a methionine biosynthetic gene encoding homoserine acetyltransferase in *Corynebacterium glutamicum*. *Mol. Cells* **8**, 286-294) - to find possible other *met* genes. It appears that *metZ* and *metA* form an operon. Expression of the genes encoding MetA and MetZ leads to overproduction of the corresponding polypeptides as can shown by gel electrophoresis.

Surprisingly, *metZ* clones can complement methiononine auxotrophic *Escherichia coli metB* mutant strains. This shows that the protein product of *metZ* catalyzes a step that can bypass the step catalyzed by the protein product of *metB*.

MetZ was also disrupted and the mutant strain showed methionine prototrophy. *Corynebacterium glutamicum* metB and metZ double mutants were also constructed. The double mutant is auxotrophic for methionine. Thus, metZ encodes a protein catalysing the reaction from O-Acetyl-Homoserine to Homocysteine, which is one step in the sulfhydrylation pathway of methionine biosynthesis. *Corynebacterium glutamicum* contains both, the transsulfuration and the sulfhydrylation pathway of methionine biosynthesis.

Introduction of *metZ* into *C. glutamicum* resulted in the expression of a 47,000 Dalton protein. Combined introduction of *metZ* and *metA* in *C. glutamicum* resulted in the appearance of *metA* and *metZ* proteins as showed by gel electrophoresis. If the Corynebacterium strain is a lysine overproducer, introduction of a plasmid containing *metZ* and *metA* resulted in a lower lysine titer but accumulation of homocysteine and methionine is detected.

In another embodiment *metZ* and *metA* were introduced into Corynebacterium glutamicum strains together with the *hom* gene, encoding the homoserine dehydrogenase, catalysing the conversion from aspartate semialdehyde to homoserine. Different *hom* genes from different organisms were chosen for this experiment. The

Corynebacterium glutamicum hom gene can be used as well as hom genes from other procaryotes like *Escherichia coli* or *Bacillus subtilis* or even the *hom* gene of eukaryotes like *Saccharomyces cerevisiae, Shizosaccharomyces pombe, Ashbya gossypii* or algae, higher plants or animals. It may be that the *hom* gene is insensitive against feed back inhibition mediated by any metabolites that occur in the biosynthetic routes of the amino acids of the aspartate familiy, like aspatrate, lysine, threonine or methionine. Such metabolites are for example aspartate, lysine, methionine, threonine, aspartyl-phosphate, aspartate semialdehyd, homoserine, cystathionine, homocysteine or any other metabolite that occurs in this biosynthetic routes. In addition to the metabolites the homoserine dehydrogenase may be insensitive against inhibition by anologes of all those metabolites or even against other compunds involved in this metabolism as there are other amino acids like cysteine or cofactors like vitamin B12 and all of its derivatives and S-adenosylmethionine and its metabolites and derivatives and anologons. The insensitivity of the homoserine dehydrogenase against all these, a part of these or only one of these compounts may either be its natural attitude or it may be the result from one or more mutations that resulted from classical mutation and selection using chemicals or irradiation or other mutagens. The mutations could also be introduced into the *hom* gene using gene technology, for example the introduction of site specific point mutations or by any method afore mentioned for the MP ore MP encoding DNA-sequences.

When a *hom* gene was combined with the *metZ* and *metA* genes and introduced into a *Corynebacterium glutamicum* strain that is a lysine overproducer, lysine accumulation was reduced and homocysteine and methionine accumulation was enhanced. A further enhancement of homocysteine and methionine concentrations can be achieved, if a lysine overproducing *Corynebacterium glutamicum* strain is used and a disruption of the *ddh* gene or the *lysA* gene was introduced prior to the transformation with DNA containing a *hom* gene and *metZ* and *metA* in combination. The overproduction of homocysteine and methionine was possible using different sulfur sources. Sulfates, thiosulfates, sulfites and also more reduced sulfur sources like H_2S and sulfides and derivatives could be used. Also organic sulfur sources like methyl mercaptan, thioglycolates, thiocyanates, thiourea, sulfur containing amino acids like cysteine and other sulfur containing compounds can be used to achieve homocysteine and methionine overproduction.

In another embodiment the *metC* gene was introduced into a *Corynebacterium glutamicum* strain using methods wich are afore mentioned. The *metC* gene can be transformed into the strain in combination with other genes like *metB*, *metA* and *metA*. Even the *hom* gene can be added. If the hom gene, the met C, metA and metB genes were combined on a vector and introduced into a *Corynebacterium glutamicum* strain homocysteine and methionine overproduction was achieved. The overproduction of homocysteine and methionine was possible using different sulfur sources. Sulfates, thiosulfates, sulfites and also more reduced sulfur sources like H_2S and sulfides and derivatives could be used. Also organic sulfur sources like methyl mercaptan, thioglycolates, thiocyanates, thiourea, sulfur containing amino acids like cysteine and other sulfur containing compounds can be used to achieve homocysteine and methionine overproduction.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patent applications, patents, published patent applications, Tables, and the sequence listing cited throughout this application are hereby incorporated by reference.

Example 1: Preparation of total genomic DNA of *Corynebacterium glutamicum* ATCC 13032

A culture of *Corynebacterium glutamicum* (ATCC 13032) was grown overnight at 30°C with vigorous shaking in BHI medium (Difco). The cells were harvested by centrifugation, the supernatant was discarded and the cells were resuspended in 5 ml buffer-I (5% of the original volume of the culture – all indicated volumes have been calculated for 100 ml of culture volume). Composition of buffer-I: 140.34 g/l sucrose, 2.46 g/l $MgSO_4 \times 7H_2O$, 10 ml/l KH_2PO_4 solution (100 g/l, adjusted to pH 6.7 with KOH), 50 ml/l M12 concentrate (10 g/l $(NH_4)_2SO_4$, 1 g/l NaCl, 2 g/l $MgSO_4 \times 7H_2O$, 0.2 g/l $CaCl_2$, 0.5 g/l yeast extract (Difco), 10 ml/l trace-elements-mix (200 mg/l $FeSO_4 \times H_2O$, 10 mg/l $ZnSO_4 \times 7 H_2O$, 3 mg/l $MnCl_2 \times 4 H_2O$, 30 mg/l H_3BO_3 20 mg/l $CoCl_2 \times 6 H_2O$, 1 mg/l $NiCl_2 \times 6 H_2O$, 3 mg/l $Na_2MoO_4 \times 2 H_2O$, 500 mg/l complexing agent (EDTA or critic acid), 100 ml/l vitamins-mix (0.2 mg/l biotin, 0.2 mg/l folic acid, 20 mg/l p-amino benzoic acid, 20 mg/l riboflavin, 40 mg/l ca-panthothenate, 140 mg/l nicotinic acid, 40 mg/l pyridoxole hydrochloride, 200 mg/l myo-inositol). Lysozyme was added to the suspension to a final concentration of 2.5 mg/ml. After an approximately 4 h incubation at 37°C, the cell wall was degraded and the resulting protoplasts are harvested by centrifugation. The pellet was washed once with 5 ml buffer-I and once with 5 ml TE-buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). The

pellet was resuspended in 4 ml TE-buffer and 0.5 ml SDS solution (10%) and 0.5 ml NaCl solution (5 M) are added. After adding of proteinase K to a final concentration of 200 µg/ml, the suspension is incubated for ca.18 h at 37°C. The DNA was purified by extraction with phenol, phenol-chloroform-isoamylalcohol and chloroform-isoamylalcohol using standard procedures. Then, the DNA was precipitated by adding 1/50 volume of 3 M sodium acetate and 2 volumes of ethanol, followed by a 30 min incubation at -20°C and a 30 min centrifugation at 12,000 rpm in a high speed centrifuge using a SS34 rotor (Sorvall). The DNA was dissolved in 1 ml TE-buffer containing 20 µg/ml RNaseA and dialysed at 4°C against 1000 ml TE-buffer for at least 3 hours. During this time, the buffer was exchanged 3 times. To aliquots of 0.4 ml of the dialysed DNA solution, 0.4 ml of 2 M LiCl and 0.8 ml of ethanol are added. After a 30 min incubation at -20°C, the DNA was collected by centrifugation (13,000 rpm, Biofuge Fresco, Heraeus, Hanau, Germany). The DNA pellet was dissolved in TE-buffer. DNA prepared by this procedure could be used for all purposes, including southern blotting or construction of genomic libraries.

Example 2: Construction of genomic libraries in *Escherichia coli* of *Corynebacterium glutamicum* ATCC13032.

Using DNA prepared as described in Example 1, cosmid and plasmid libraries were constructed according to known and well established methods (*see e.g.*, Sambrook, J. *et al.* (1989) "Molecular Cloning : A Laboratory Manual", Cold Spring Harbor Laboratory Press, or Ausubel, F.M. *et al.* (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.)

Any plasmid or cosmid could be used. Of particular use were the plasmids pBR322 (Sutcliffe, J.G. (1979) *Proc. Natl. Acad. Sci. USA*, 75:3737-3741); pACYC177 (Change & Cohen (1978) *J. Bacteriol* 134:1141-1156), plasmids of the pBS series (pBSSK+, pBSSK- and others; Stratagene, LaJolla, USA), or cosmids as SuperCos1 (Stratagene, LaJolla, USA) or Lorist6 (Gibson, T.J., Rosenthal A. and Waterson, R.H. (1987) *Gene* 53:283-286. Gene libraries specifically for use in *C. glutamicum* may be constructed using plasmid pSL109 (Lee, H.-S. and A. J. Sinskey (1994) *J. Microbiol. Biotechnol.* 4: 256-263).

Example 3: DNA Sequencing and Computational Functional Analysis

Genomic libraries as described in Example 2 were used for DNA sequencing according to standard methods, in particular by the chain termination method using ABI377 sequencing machines (see *e.g.*, Fleischman, R.D. *et al.* (1995) "Whole-genome Random

Sequencing and Assembly of Haemophilus Influenzae Rd., *Science*, 269:496-512). Sequencing primers with the following nucleotide sequences were used: 5'-GGAAACAGTATGACCATG-3' or 5'-GTAAAACGACGGCCAGT-3'.

Example 4: *In vivo* Mutagenesis

In vivo mutagenesis of *Corynebacterium glutamicum* can be performed by passage of plasmid (or other vector) DNA through *E. coli* or other microorganisms (*e.g. Bacillus* spp. or yeasts such as *Saccharomyces cerevisiae*) which are impaired in their capabilities to maintain the integrity of their genetic information. Typical mutator strains have mutations in the genes for the DNA repair system (*e.g.*, mutHLS, mutD, mutT, etc.; for reference, see Rupp, W.D. (1996) DNA repair mechanisms, in: *Escherichia coli* and *Salmonella*, p. 2277-2294, ASM: Washington.) Such strains are well known to those of ordinary skill in the art. The use of such strains is illustrated, for example, in Greener, A. and Callahan, M. (1994) *Strategies* 7: 32-34.

Example 5: DNA Transfer Between *Escherichia coli* and *Corynebacterium glutamicum*

Several *Corynebacterium* and *Brevibacterium* species contain endogenous plasmids (as *e.g.*, pHM1519 or pBL1) which replicate autonomously (for review see, *e.g.*, Martin, J.F. *et al.* (1987) *Biotechnology*, 5:137-146). Shuttle vectors for *Escherichia coli* and *Corynebacterium glutamicum* can be readily constructed by using standard vectors for *E. coli* (Sambrook, J. *et al.* (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press or Ausubel, F.M. *et al.* (1994) "Current Protocols in Molecular Biology", John Wiley & Sons) to which a origin or replication for and a suitable marker from *Corynebacterium glutamicum* is added. Such origins of replication are preferably taken from endogenous plasmids isolated from *Corynebacterium* and *Brevibacterium* species. Of particular use as transformation markers for these species are genes for kanamycin resistance (such as those derived from the Tn5 or Tn903 transposons) or chloramphenicol (Winnacker, E.L. (1987) "From Genes to Clones — Introduction to Gene Technology, VCH, Weinheim). There are numerous examples in the literature of the construction of a wide variety of shuttle vectors which replicate in both *E. coli and C. glutamicum*, and which can be used for several purposes, including gene over-expression (for reference, see *e.g.*, Yoshihama, M. *et al.* (1985) *J. Bacteriol.* 162:591-597,

Martin J.F. *et al.* (1987) *Biotechnology*, 5:137-146 and Eikmanns, B.J. *et al.* (1991) *Gene*, 102:93-98).

Using standard methods, it is possible to clone a gene of interest into one of the shuttle vectors described above and to introduce such a hybrid vectors into strains of *Corynebacterium glutamicum*. Transformation of *C. glutamicum* can be achieved by protoplast transformation (Kastsumata, R. *et al.* (1984) *J. Bacteriol.* 159306-311), electroporation (Liebl, E. *et al.* (1989) *FEMS Microbiol. Letters*, 53:399-303) and in cases where special vectors are used, also by conjugation (as described *e.g.* in Schäfer, A *et al.* (1990) *J. Bacteriol.* 172:1663-1666). It is also possible to transfer the shuttle vectors for *C. glutamicum* to *E. coli* by preparing plasmid DNA from *C. glutamicum* (using standard methods well-known in the art) and transforming it into *E. coli*. This transformation step can be performed using standard methods, but it is advantageous to use an Mcr-deficient *E. coli* strain, such as NM522 (Gough & Murray (1983) *J. Mol. Biol.* 166:1-19).

Genes may be overexpressed in *C. glutamicum* strains using plasmids which comprise pCG1 (U.S. Patent No. 4,617,267) or fragments thereof, and optionally the gene for kanamycin resistance from TN903 (Grindley, N.D. and Joyce, C.M. (1980) *Proc. Natl. Acad. Sci. USA* 77(12): 7176-7180). In addition, genes may be overexpressed in *C. glutamicum* strains using plasmid pSL109 (Lee, H.-S. and A. J. Sinskey (1994) *J. Microbiol. Biotechnol.* 4: 256-263).

Aside from the use of replicative plasmids, gene overexpression can also be achieved by integration into the genome. Genomic integration in *C. glutamicum* or other Corynebacterium or Brevibacterium species may be accomplished by well-known methods, such as homologous recombination with genomic region(s), restriction endonuclease mediated integration (REMI) (see, *e.g.*, DE Patent 19823834), or through the use of transposons. It is also possible to modulate the activity of a gene of interest by modifying the regulatory regions (*e.g.*, a promoter, a repressor, and/or an enhancer) by sequence modification, insertion, or deletion using site-directed methods (such as homologous recombination) or methods based on random events (such as transposon mutagenesis or REMI). Nucleic acid sequences which function as transcriptional terminators may also be inserted 3' to the coding region of one or more genes of the invention; such terminators are well-known in the art and are described, for example, in Winnacker, E.L. (1987) From Genes to Clones - Introduction to Gene Technology. VCH: Weinheim.

Example 6: Assessment of the Expression of the Mutant Protein

Observations of the activity of a mutated protein in a transformed host cell rely on the fact that the mutant protein is expressed in a similar fashion and in a similar quantity to that of the wild-type protein. A useful method to ascertain the level of transcription of the mutant gene (an indicator of the amount of mRNA available for translation to the gene product) is to perform a Northern blot (for reference see, for example, Ausubel *et al.* (1988) Current Protocols in Molecular Biology, Wiley: New York), in which a primer designed to bind to the gene of interest is labeled with a detectable tag (usually radioactive or chemiluminescent), such that when the total RNA of a culture of the organism is extracted, run on gel, transferred to a stable matrix and incubated with this probe, the binding and quantity of binding of the probe indicates the presence and also the quantity of mRNA for this gene. This information is evidence of the degree of transcription of the mutant gene. Total cellular RNA can be prepared from *Corynebacterium glutamicum* by several methods, all well-known in the art, such as that described in Bormann, E.R. *et al.* (1992) *Mol. Microbiol.* 6: 317-326.

To assess the presence or relative quantity of protein translated from this mRNA, standard techniques, such as a SDS-Polyacrylamide Gelelectrophoresis and Western blot, may be employed (see, for example, Ausubel *et al.* (1988) Current Protocols in Molecular Biology, Wiley: New York). In this process, total cellular proteins are extracted, separated by gel electrophoresis, transferred to a matrix such as nitrocellulose, and incubated with a probe, such as an antibody, which specifically binds to the desired protein. This probe is generally tagged with a chemiluminescent or colorimetric label which may be readily detected. The presence and quantity of label observed indicates the presence and quantity of the desired mutant protein present in the cell.

Example 7: Growth of Escherichia coli and Genetically Modified Corynebacterium glutamicum – Media and Culture Conditions

E. coli strains are routinely grown in MB and LB broth, respectively (Follettie, M.T., Peoples, O., Agoropoulou, C., and Sinskey, A J. (1993) Gene structure and expression of the Corynebacterium flavum N13 ask-asd operon. J. Bacteriol. **175**, 4096-4103). Minimal media for E. coli is M9 and modified MCGC (Yoshihama, M., Higashiro, K., Rao, E.A., Akedo, M., Shanabruch, W G., Follettie, M.T., Walker, G.C., and Sinskey, A.J.

(1985) Cloning vector system for Corynebacterium glutamicum. J. Bacteriol. 162, 591-507), respectively. Glucose was added a final concentration of 1%. Antibiotics were added in the following amounts (micrograms per milliliter): ampicillin, 50; kanamycin, 25; nalidixic acid, 25. Amino acids, vitamins, and other supplements were added in the following amounts: methionine, 9.3 mM; arginine, 9.3 mM; histidine, 9.3 mM; thiamine, 0.05 mM. E. coli cells were routinely grown at 37°C, respectively.

Genetically modified *Corynebacteria* are cultured in synthetic or natural growth media. A number of different growth media for Corynebacteria are both well-known and readily available (Lieb *et al.* (1989) *Appl. Microbiol. Biotechnol.*, 32:205-210; von der Osten *et al.* (1998) Biotechnology Letters, 11:11-16; Patent DE 4,120,867; Liebl (1992) "The Genus *Corynebacterium*, in: The Procaryotes, Volume II, Balows, A. *et al.*, eds. Springer-Verlag). These media consist of one or more carbon sources, nitrogen sources, inorganic salts, vitamins and trace elements. Preferred carbon sources are sugars, such as mono-, di-, or polysaccharides. For example, glucose, fructose, mannose, galactose, ribose, sorbose, ribulose, lactose, maltose, sucrose, raffinose, starch or cellulose serve as very good carbon sources. It is also possible to supply sugar to the media via complex compounds such as molasses or other by-products from sugar refinement. It can also be advantageous to supply mixtures of different carbon sources. Other possible carbon sources are alcohols and organic acids, such as methanol, ethanol, acetic acid or lactic acid. Nitrogen sources are usually organic or inorganic nitrogen compounds, or materials which contain these compounds. Exemplary nitrogen sources include ammonia gas or ammonia salts, such as NH_4Cl or $(NH_4)_2SO_4$, NH_4OH, nitrates, urea, amino acids or complex nitrogen sources like corn steep liquor, soy bean flour, soy bean protein, yeast extract, meat extract and others.

The overproduction of sulfur containig amino acids like homocysteine and methionine was possible using different sulfur sources. Sulfates, thiosulfates, sulfites and also more reduced sulfur sources like H_2S and sulfides and derivatives can be used. Also organic sulfur sources like methyl mercaptan, thioglycolates, thiocyanates, thiourea, sulfur containing amino acids like cysteine and other sulfur containing compounds can be used to achieve homocysteine and methionine overproduction.

Inorganic salt compounds which may be included in the media include the chloride-, phosphorous- or sulfate- salts of calcium, magnesium, sodium, cobalt, molybdenum, potassium, manganese, zinc, copper and iron. Chelating compounds can be added to the medium to keep the metal ions in solution. Particularly useful chelating compounds include dihydroxyphenols, like catechol or protocatechuate, or organic acids, such as citric acid. It is typical for the media to also contain other growth factors, such as vitamins or growth promoters, examples of which include biotin, riboflavin, thiamin, folic acid, nicotinic acid, pantothenate and pyridoxin. Growth factors and salts frequently originate from complex media components such as yeast extract, molasses, corn steep liquor and others. The exact composition of the media compounds depends strongly on the immediate experiment and is individually decided for each specific case. Information about media optimization is available in the textbook "Applied Microbiol. Physiology, A Practical Approach (*eds.* P.M. Rhodes, P.F. Stanbury, IRL Press (1997) pp. 53-73, ISBN 0 19 963577 3). It is also possible to select growth media from commercial suppliers, like standard 1 (Merck) or BHI (grain heart infusion, DIFCO) or others.

All medium components are sterilized, either by heat (20 minutes at 1.5 bar and 121°C) or by sterile filtration. The components can either be sterilized together or, if necessary, separately. All media components can be present at the beginning of growth, or they can optionally be added continuously or batchwise.

Culture conditions are defined separately for each experiment. The temperature should be in a range between 15°C and 45°C. The temperature can be kept constant or can be altered during the experiment. The pH of the medium should be in the range of 5 to 8.5, preferably around 7.0, and can be maintained by the addition of buffers to the media. An exemplary buffer for this purpose is a potassium phosphate buffer. Synthetic buffers such as MOPS, HEPES, ACES and others can alternatively or simultaneously be used. It is also possible to maintain a constant culture pH through the addition of NaOH or NH_4OH during growth. If complex medium components such as yeast extract are utilized, the necessity for additional buffers may be reduced, due to the fact that many complex compounds have high buffer capacities. If a fermentor is utilized for culturing the micro-organisms, the pH can also be controlled using gaseous ammonia.

The incubation time is usually in a range from several hours to several days. This time is selected in order to permit the maximal amount of product to accumulate in the broth. The

disclosed growth experiments can be carried out in a variety of vessels, such as microtiter plates, glass tubes, glass flasks or glass or metal fermentors of different sizes. For screening a large number of clones, the microorganisms should be cultured in microtiter plates, glass tubes or shake flasks, either with or without baffles. Preferably 100 ml shake flasks are used, filled with 10% (by volume) of the required growth medium. The flasks should be shaken on a rotary shaker (amplitude 25 mm) using a speed-range of 100-300 rpm. Evaporation losses can be diminished by the maintenance of a humid atmosphere; alternatively, a mathematical correction for evaporation losses should be performed.

If genetically modified clones are tested, an unmodified control clone or a control clone containing the basic plasmid without any insert should also be tested. The medium is inoculated to an OD_{600} of 0.5 - 1.5 using cells grown on agar plates, such as CM plates (10 g/l glucose, 2,5 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract, 22 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract, 22 g/l agar, pH 6.8 with 2M NaOH) that had been incubated at 30°C. Inoculation of the media is accomplished by either introduction of a saline suspension of *C. glutamicum* cells from CM plates or addition of a liquid preculture of this bacterium.

Example 8: *In vitro* Analysis of the Function of Mutant Proteins

The determination of activities and kinetic parameters of enzymes is well established in the art. Experiments to determine the activity of any given altered enzyme must be tailored to the specific activity of the wild-type enzyme, which is well within the ability of one of ordinary skill in the art. Overviews about enzymes in general, as well as specific details concerning structure, kinetics, principles, methods, applications and examples for the determination of many enzyme activities may be found, for example, in the following references: Dixon, M., and Webb, E.C., (1979) Enzymes. Longmans: London; Fersht, (1985) Enzyme Structure and Mechanism. Freeman: New York; Walsh, (1979) Enzymatic Reaction Mechanisms. Freeman: San Francisco; Price, N.C., Stevens, L. (1982) Fundamentals of Enzymology. Oxford Univ. Press: Oxford; Boyer, P.D., ed. (1983) The Enzymes, 3rd ed. Academic Press: New York; Bisswanger, H., (1994) Enzymkinetik, 2nd ed. VCH: Weinheim (ISBN 3527300325); Bergmeyer, H.U., Bergmeyer, J., Graßl, M., eds. (1983-1986) Methods of Enzymatic Analysis, 3rd ed., vol. I-XII, Verlag Chemie: Weinheim; and

Ullmann's Encyclopedia of Industrial Chemistry (1987) vol. A9, "Enzymes". VCH: Weinheim, p. 352-363.

The activity of proteins which bind to DNA can be measured by several well-established methods, such as DNA band-shift assays (also called gel retardation assays). The effect of such proteins on the expression of other molecules can be measured using reporter gene assays (such as that described in Kolmar, H. *et al.* (1995) *EMBO J.* 14: 3895-3904 and references cited therein). Reporter gene test systems are well known and established for applications in both pro- and eukaryotic cells, using enzymes such as beta-galactosidase, green fluorescent protein, and several others.

The determination of activity of membrane-transport proteins can be performed according to techniques such as those described in Gennis, R.B. (1989) "Pores, Channels and Transporters", in Biomembranes, Molecular Structure and Function, Springer: Heidelberg, p. 85-137; 199-234; and 270-322.

Example 9: Analysis of Impact of Mutant Protein on the Production of the Desired Product

The effect of the genetic modification in *C. glutamicum* on production of a desired compound (such as an amino acid) can be assessed by growing the modified microorganism under suitable conditions (such as those described above) and analyzing the medium and/or the cellular component for increased production of the desired product (*i.e.*, an amino acid). Such analysis techniques are well known to one of ordinary skill in the art, and include spectroscopy, thin layer chromatography, staining methods of various kinds, enzymatic and microbiological methods, and analytical chromatography such as high performance liquid chromatography (see, for example, Ullman, Encyclopedia of Industrial Chemistry, vol. A2, p. 89-90 and p. 443-613, VCH: Weinheim (1985); Fallon, A. *et al.*, (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17; Rehm *et al.* (1993) Biotechnology, vol. 3, Chapter III: "Product recovery and purification", page 469-714, VCH: Weinheim; Belter, P.A. *et al.* (1988) Bioseparations: downstream processing for biotechnology, John Wiley and Sons; Kennedy, J.F. and Cabral, J.M.S. (1992) Recovery processes for biological materials, John Wiley and Sons; Shaeiwitz, J.A. and Henry, J.D. (1988) Biochemical separations, in: Ulmann's Encyclopedia of Industrial Chemistry, vol. B3, Chapter 11, page 1-27, VCH: Weinheim; and Dechow, F.J. (1989)

Separation and purification techniques in biotechnology, Noyes Publications.)

In addition to the measurement of the final product of fermentation, it is also possible to analyze other components of the metabolic pathways utilized for the production of the desired compound, such as intermediates and side-products, to determine the overall efficiency of production of the compound. Analysis methods include measurements of nutrient levels in the medium (*e.g.*, sugars, hydrocarbons, nitrogen sources, phosphate, and other ions), measurements of biomass composition and growth, analysis of the production of common metabolites of biosynthetic pathways, and measurement of gasses produced during fermentation. Standard methods for these measurements are outlined in Applied Microbial Physiology, A Practical Approach, P.M. Rhodes and P.F. Stanbury, eds., IRL Press, p. 103-129; 131-163; and 165-192 (ISBN: 0199635773) and references cited therein.

Example 10: Purification of the Desired Product from *C. glutamicum* Culture

Recovery of the desired product from the *C. glutamicum* cells or supernatant of the above-described culture can be performed by various methods well known in the art. If the desired product is not secreted from the cells, the cells can be harvested from the culture by low-speed centrifugation, the cells can be lysed by standard techniques, such as mechanical force or sonication. The cellular debris is removed by centrifugation, and the supernatant fraction containing the soluble proteins is retained for further purification of the desired compound. If the product is secreted from the *C. glutamicum* cells, then the cells are removed from the culture by low-speed centrifugation, and the supernate fraction is retained for further purification.

The supernatant fraction from either purification method is subjected to chromatography with a suitable resin, in which the desired molecule is either retained on a chromatography resin while many of the impurities in the sample are not, or where the impurities are retained by the resin while the sample is not. Such chromatography steps may be repeated as necessary, using the same or different chromatography resins. One of ordinary skill in the art would be well-versed in the selection of appropriate chromatography resins and in their most efficacious application for a particular molecule to be purified. The purified product may be concentrated by filtration or ultrafiltration, and stored

at a temperature at which the stability of the product is maximized.

There are a wide array of purification methods known to the art and the preceding method of purification is not meant to be limiting. Such purification techniques are described, for example, in Bailey, J.E. & Ollis, D.F. Biochemical Engineering Fundamentals, McGraw-Hill: New York (1986).

The identity and purity of the isolated compounds may be assessed by techniques standard in the art. These include high-performance liquid chromatography (HPLC), spectroscopic methods, staining methods, thin layer chromatography, NIRS, enzymatic assay, or microbiologically. Such analysis methods are reviewed in: Patek *et al.* (1994) *Appl. Environ. Microbiol.* 60: 133-140; Malakhova *et al.* (1996) *Biotekhnologiya* 11: 27-32; and Schmidt *et al.* (1998) *Bioprocess Engineer.* 19: 67-70. Ulmann's Encyclopedia of Industrial Chemistry, (1996) vol. A27, VCH: Weinheim, p. 89-90, p. 521-540, p. 540-547, p. 559-566, 575-581 and p. 581-587; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A. *et al.* (1987) Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17.

Example 11: Analysis of the Gene Sequences of the Invention

The comparison of sequences and determination of percent homology between two sequences are art-known techniques, and can be accomplished using a mathematical algorithm, such as the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci.* USA 87:2264-68, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci.* USA 90:5873-77. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to MP nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to MP protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, one of ordinary skill in the art will know how to optimize the parameters of the program (*e.g.*, XBLAST and NBLAST) for the specific sequence being analyzed.

Another example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Meyers and Miller ((1988) *Comput. Appl. Biosci.* 4: 11-17). Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Additional algorithms for sequence analysis are known in the art, and include ADVANCE and ADAM. described in Torelli and Robotti (1994) *Comput. Appl. Biosci.* 10:3-5; and FASTA, described in Pearson and Lipman (1988) *P.N.A.S.* 85:2444-8.

The percent homology between two amino acid sequences can also be accomplished using the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 12, 10, 8, 6, or 4 and a length weight of 2, 3, or 4. The percent homology between two nucleic acid sequences can be accomplished using the GAP program in the GCG software package, using standard parameters, such as a gap weight of 50 and a length weight of 3.

A comparative analysis of the gene sequences of the invention with those present in Genbank has been performed using techniques known in the art (see, *e.g.*, Bexevanis and Ouellette, eds. (1998) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. John Wiley and Sons: New York).

The gene sequences of the invention were compared on basis of their amino acid sequences to known genes by using the program CLUSTAL (Higgins et al. (1996) Using CLUSTAL for multiple sequence alignments, Methods in Enzymology 266, 383-402) using the standard parameters (PAIRWISE ALIGNMENT PARAMETERS: Gap penalty= 3, K-tuple (word) size= 1, No. of top diagonals= 5, Window size= 5; MULTIPLE ALIGNMENT PARAMETERS: Gap Opening Penalty= 10.00, Gap Extension Penalty= 0.05, Protein weight matrix= PAM250). Homology between two sequences is the function of the number of identical positions in all sequences (i.e. % homology = number of identical positions/total number of positions x 100). The results of this analysis are set forth in Table 3.

Example 12: Construction and Operation of DNA Microarrays

The sequences of the invention may additionally be used in the construction and application of DNA microarrays (the design, methodology, and uses of DNA arrays are well known in the art,

and are described, for example, in Schena, M. *et al.* (1995) *Science* 270: 467-470; Wodicka, L. *et al.* (1997) *Nature Biotechnology* 15: 1359-1367; DeSaizieu, A. *et al.* (1998) *Nature Biotechnology* 16: 45-48; and DeRisi, J.L. *et al.* (1997) *Science* 278: 680-686).

DNA microarrays are solid or flexible supports consisting of nitrocellulose, nylon, glass, silicone, or other materials. Nucleic acid molecules may be attached to the surface in an ordered manner. After appropriate labeling, other nucleic acids or nucleic acid mixtures can be hybridized to the immobilized nucleic acid molecules, and the label may be used to monitor and measure the individual signal intensities of the hybridized molecules at defined regions. This methodology allows the simultaneous quantification of the relative or absolute amount of all or selected nucleic acids in the applied nucleic acid sample or mixture. DNA microarrays, therefore, permit an analysis of the expression of multiple (as many as 6800 or more) nucleic acids in parallel (see, *e.g.*, Schena, M. (1996) *BioEssays* 18(5): 427-431).

The sequences of the invention may be used to design oligonucleotide primers which are able to amplify defined regions of one or more *C. glutamicum* genes by a nucleic acid amplification reaction such as the polymerase chain reaction. The choice and design of the 5' or 3' oligonucleotide primers or of appropriate linkers allows the covalent attachment of the resulting PCR products to the surface of a support medium described above (and also described, for example, Schena, M. *et al.* (1995) *Science* 270: 467-470).

Nucleic acid microarrays may also be constructed by *in situ* oligonucleotide synthesis as described by Wodicka, L. *et al.* (1997) *Nature Biotechnology* 15: 1359-1367. By photolithographic methods, precisely defined regions of the matrix are exposed to light. Protective groups which are photolabile are thereby activated and undergo nucleotide addition, whereas regions that are masked from light do not undergo any modification. Subsequent cycles of protection and light activation permit the synthesis of different oligonucleotides at defined positions. Small, defined regions of the genes of the invention may be synthesized on microarrays by solid phase oligonucleotide synthesis.

The nucleic acid molecules of the invention present in a sample or mixture of nucleotides may be hybridized to the microarrays. These nucleic acid molecules can be labeled according to standard methods. In brief, nucleic acid molecules (*e.g.*, mRNA molecules or DNA molecules) are labeled by the incorporation of

isotopically or fluorescently labeled nucleotides, *e.g.*, during reverse transcription or DNA synthesis. Hybridization of labeled nucleic acids to microarrays is described (*e.g.*, in Schena, M. *et al.* (1995) *supra*; Wodicka, L. *et al.* (1997), *supra*; and DeSaizieu A. *et al.* (1998), *supra*). The detection and quantification of the hybridized molecule are tailored to the specific incorporated label. Radioactive labels can be detected, for example, as described in Schena, M. *et al.* (1995) *supra*) and fluorescent labels may be detected, for example, by the method of Shalon *et al.* (1996) *Genome Research* 6: 639-645).

The application of the sequences of the invention to DNA microarray technology, as described above, permits comparative analyses of different strains of *C. glutamicum* or other Corynebacteria. For example, studies of inter-strain variations based on individual transcript profiles and the identification of genes that are important for specific and/or desired strain properties such as pathogenicity, productivity and stress tolerance are facilitated by nucleic acid array methodologies. Also, comparisons of the profile of expression of genes of the invention during the course of a fermentation reaction are possible using nucleic acid array technology.

Example 13: Analysis of the Dynamics of Cellular Protein Populations (Proteomics)

The genes, compositions, and methods of the invention may be applied to study the interactions and dynamics of populations of proteins, termed 'proteomics'. Protein populations of interest include, but are not limited to, the total protein population of *C. glutamicum* (*e.g.*, in comparison with the protein populations of other organisms), those proteins which are active under specific environmental or metabolic conditions (*e.g.*, during fermentation, at high or low temperature, or at high or low pH), or those proteins which are active during specific phases of growth and development.

Protein populations can be analyzed by various well-known techniques, such as gel electrophoresis. Cellular proteins may be obtained, for example, by lysis or extraction, and may be separated from one another using a variety of electrophoretic techniques. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) separates proteins largely on the basis of their molecular weight. Isoelectric focusing polyacrylamide gel electrophoresis (IEF-PAGE) separates proteins by their isoelectric point (which reflects not only the amino acid sequence but also posttranslational modifications of the

protein). Another, more preferred method of protein analysis is the consecutive combination of both IEF-PAGE and SDS-PAGE, known as 2-D-gel electrophoresis (described, for example, in Hermann *et al.* (1998) *Electrophoresis* 19: 3217-3221; Fountoulakis *et al.* (1998) *Electrophoresis* 19: 1193-1202; Langen *et al.* (1997) *Electrophoresis* 18: 1184-1192; Antelmann *et al.* (1997) *Electrophoresis* 18: 1451-1463). Other separation techniques may also be utilized for protein separation, such as capillary gel electrophoresis; such techniques are well known in the art.

Proteins separated by these methodologies can be visualized by standard techniques, such as by staining or labeling. Suitable stains are known in the art, and include Coomassie Brilliant Blue, silver stain, or fluorescent dyes such as Sypro Ruby (Molecular Probes). The inclusion of radioactively labeled amino acids or other protein precursors (*e.g.*, ^{35}S-methionine, ^{35}S-cysteine, ^{14}C-labelled amino acids, ^{15}N-amino acids, $^{15}NO_3$ or $^{15}NH_4^+$ or ^{13}C-labelled amino acids) in the medium of *C. glutamicum* permits the labeling of proteins from these cells prior to their separation. Similarly, fluorescent labels may be employed. These labeled proteins can be extracted, isolated and separated according to the previously described techniques.

Proteins visualized by these techniques can be further analyzed by measuring the amount of dye or label used. The amount of a given protein can be determined quantitatively using, for example, optical methods and can be compared to the amount of other proteins in the same gel or in other gels. Comparisons of proteins on gels can be made, for example, by optical comparison, by spectroscopy, by image scanning and analysis of gels, or through the use of photographic films and screens. Such techniques are well-known in the art.

To determine the identity of any given protein, direct sequencing or other standard techniques may be employed. For example, N- and/or C-terminal amino acid sequencing (such as Edman degradation) may be used, as may mass spectrometry (in particular MALDI or ESI techniques (see, *e.g.*, Langen *et al.* (1997) *Electrophoresis* 18: 1184-1192)). The protein sequences provided herein can be used for the identification of *C. glutamicum* proteins by these techniques.

The information obtained by these methods can be used to compare patterns of protein presence, activity, or modification between different samples from various biological conditions (*e.g.*, different organisms, time points of fermentation, media conditions, or different biotopes, among others). Data obtained

from such experiments alone, or in combination with other techniques, can be used for various applications, such as to compare the behavior of various organisms in a given (*e.g.*, metabolic) situation, to increase the productivity of strains which produce fine chemicals or to increase the efficiency of the production of fine chemicals.

Equivalents

Those of ordinary skill in the art will recognize, or will be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

Claims

1. An isolated *Corynebacterium glutamicum* nucleic acid molecule selected from the group consisting of those sequences set forth as odd-numbered SEQ ID NOs of the Sequence Listing, or a portion thereof, as set forth in Table 1.

2. An isolated nucleic acid molecule which encodes a polypeptide sequence selected from the group consisting of those sequences set forth as even-numbered SEQ ID NOs of the Sequence Listing, as set forth in Table 1.

3. An isolated nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide selected from the group of amino acid sequences consisting of those sequences set forth as even-numbered SEQ ID NOs of the Sequence Listing, as set forth in Table 1.

4. An isolated nucleic acid molecule comprising a nucleotide sequence which is at least 63% homologous on basis of its amino acid sequence to a nucleotide sequence selected from the group consisting of those sequences which encode for an amino acid sequence as set forth as SEQ ID NO 2 of the Sequence Listing, or a portion thereof, or sequence which is at least 71% homologous on basis of its amino acid sequence to a nucleotide sequence selected from the group consisting of those sequences which encode for an amino acid sequence as set forth as SEQ ID NO 4 of the Sequence Listing, or a portion thereof.

5. An isolated nucleic acid molecule comprising a fragment of at least 15 nucleotides of a nucleic acid comprising a nucleotide sequence selected from the group consisting of those sequences set forth as odd-numbered SEQ ID NOs of the Sequence Listing, as set forth in Table 1.

6. An isolated nucleic acid molecule which hybridizes to the nucleic acid molecule of any one of claims 1-5 under stringent conditions.

7. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1-6 or a portion thereof and a nucleotide sequence encoding a heterologous polypeptide.

8. A DNA-construct comprising the nucleic acid molecule of any one of claims 1-7 and a regulatory sequence.

9. A vector comprising the nucleic acid molecule of any one of claims 1-7.

10. A vector of claim 9 comprising in addition one ore more copies of the same or different nucleic acid molecule of table 4 provided the nucleic acid molecule pertains methionine or of table 5 provided the nucleic acid molecule pertains trehalose.

11. The vector of any one of the claims 9 or 10, which is an expression vector.

12. A host cell transfected with the expression vector of claim 11.

13. The host cell of claim 12, wherein said cell is a micro-organism.

14. The host cell of claim 13, wherein said cell belongs to the genus *Corynebacterium* or *Brevibacterium*.

15. The host cell of claim 12, wherein the expression of said nucleic acid molecule results in the modulation in production of a fine chemical from said cell.

16. The host cell of claim 15, wherein said fine chemical is selected from the group consisting of: organic acids, non-proteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, polyketides, and enzymes.

17. A method of producing a polypeptide comprising culturing the host cell of claim 12 in an appropriate culture medium to, thereby, produce the polypeptide.

18. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of those sequences set forth as even-numbered SEQ ID NOs of the Sequence Listing, as set forth in Table 1.

19. An isolated polypeptide comprising a naturally occurring allelic variant of a polypeptide comprising an amino acid sequence selected from the group consisting of those sequences set forth as even-numbered SEQ ID NOs of the Sequence Listing, or a portion thereof, as set forth in Table 1.

20. The isolated polypeptide of any of claims 18 or 19, further comprising heterologous amino acid sequences.

21. An isolated polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 63% homologous to a nucleic acid selected from the group consisting of those sequences set forth as odd-numbered SEQ ID NOs of the Sequence Listing, as set forth in Table 1.

22. An isolated polypeptide comprising an amino acid sequence which is at least 63% homologous to an amino acid sequence selected from the group consisting of those sequences set forth as even-numbered SEQ ID NOs of the Sequence Listing, as set forth in Table 1.

23. A method for producing a fine chemical, comprising culturing a cell containing a vector of claim 11 such that the fine chemical is produced.

24. The method of claim 23, wherein said method further comprises the step of recovering the fine chemical from said culture.

25. The method of claim 23, wherein said method further comprises the step of transfecting said cell with the vector of claim 11 to result in a cell containing said vector.

26. The method of claim 23, wherein said cell belongs to the genus *Corynebacterium* or *Brevibacterium*.

27. The method of claim 23, wherein said cell is selected from the group consisting of: *Corynebacterium glutamicum, Corynebacterium herculis, Corynebacterium, lilium, Corynebacterium acetoacidophilum, Corynebacterium acetoglutamicum, Corynebacterium acetophilum, Corynebacterium ammoniagenes, Corynebacterium fujiokense, Corynebacterium nitrilophilus, Brevibacterium ammoniagenes, Brevibacterium butanicum, Brevibacterium divaricatum, Brevibacterium flavum, Brevibacterium healii, Brevibacterium ketoglutamicum, Brevibacterium ketosoreductum, Brevibacterium lactofermentum, Brevibacterium*

linens, Brevibacterium paraffinolyticum, and those strains set forth in Table 2.

28. The method of claim 23, wherein expression of the nucleic acid molecule from said vector results in modulation of production of said fine chemical.

29. The method of claim 23, wherein said fine chemical is selected from the group consisting of: organic acids, non-proteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, polyketides, and enzymes.

30. The method of claim 23, wherein said fine chemical is an amino acid or a carbohydrate.

31. The method of claim 30, wherein said amino acid carbohydrate is drawn from the group consisting of: methionine or trehalose.

32. A method for producing a fine chemical, comprising culturing a cell whose genomic DNA has been altered by the inclusion of a nucleic acid molecule of any one of claims 1-7.

33. A method for producing a fine chemical of claim 32 comprising in addition one ore more copies of the same or different nucleic acid molecule of table 4 provided the nucleic acid molecule pertains methionine or of table 5 provided the nucleic acid molecule pertains trehalose.

34. A method for diagnosing the presence or activity of *Corynebacterium diphtheriae* in a subject, comprising detecting the presence of one or more of SEQ ID NOs 1 through 4 of the Sequence Listing in the subject, thereby diagnosing the presence or activity of *Corynebacterium diphtheriae* in the subject.

35. A host cell comprising a nucleic acid molecule selected from the group consisting of the nucleic acid molecules set forth as odd-numbered SEQ ID NOs of the Sequence Listing, wherein the nucleic acid molecule is disrupted.

36. A host cell comprising a nucleic acid molecule selected from the group consisting of the nucleic acid molecules set forth as odd-numbered SEQ ID NOs in the Sequence Listing, wherein the nucleic acid molecule comprises one or more nucleic acid modifications from the sequence set forth as odd-numbered SEQ ID NOs of the Sequence Listing.

37. A host cell comprising a nucleic acid molecule selected from the group consisting of the nucleic acid molecules set forth as odd-numbered SEQ ID NOs of the Sequence Listing, wherein the regulatory region of the nucleic acid molecule is modified relative to the wild-type regulatory region of the molecule.

SEQUENZPROTOKOLL

<110> BASF Aktiengesellschaft

<120> Novel genes of Corynebacterium

<130> 936_2000

<140> 936_2000
<141> 2000-12-22

<160> 46

<170> PatentIn Vers. 2.0

<210> 1
<211> 3793
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(3763)

<400> 1
agactagtgg cgctttgcct gtgttgctta ggcggcgttg aaaatgaact acgaatgaaa 60

agttcgggaa ttgtctaatc cgtactaagc tgtctacaca atg tct act tca gtt 115
Met Ser Thr Ser Val
1 5

act tca cca gcc cac aac aac gca cat tcc tcc gaa ttt ttg gat gcg 163
Thr Ser Pro Ala His Asn Asn Ala His Ser Ser Glu Phe Leu Asp Ala
10 15 20

ttg gca aac cat gtg ttg atc ggc gac ggc gcc atg ggc acc cag ctc 211
Leu Ala Asn His Val Leu Ile Gly Asp Gly Ala Met Gly Thr Gln Leu
25 30 35

caa ggc ttt gac ctg gac gtg gaa aag gat ttc ctt gat ctg gag ggg 259
Gln Gly Phe Asp Leu Asp Val Glu Lys Asp Phe Leu Asp Leu Glu Gly
40 45 50

tgt aat gag att ctc aac gac acc cgc cct gat gtg ttg agg cag att 307
Cys Asn Glu Ile Leu Asn Asp Thr Arg Pro Asp Val Leu Arg Gln Ile
55 60 65

cac cgc gcc tac ttt gag gcg gga gct gac ttg gtt gag acc aat act 355
His Arg Ala Tyr Phe Glu Ala Gly Ala Asp Leu Val Glu Thr Asn Thr
70 75 80 85

ttt ggt tgc aac ctg ccg aac ttg gcg gat tat gac atc gct gat cgt 403
Phe Gly Cys Asn Leu Pro Asn Leu Ala Asp Tyr Asp Ile Ala Asp Arg
90 95 100

```
tgc cgt gag ctt gcc tac aag ggc act gca gtg gct agg gaa gtg gct   451
Cys Arg Glu Leu Ala Tyr Lys Gly Thr Ala Val Ala Arg Glu Val Ala
            105                 110                 115

gat gag atg ggg ccg ggc cga aac ggc atg cgg cgt ttc gtg gtt ggt   499
Asp Glu Met Gly Pro Gly Arg Asn Gly Met Arg Arg Phe Val Val Gly
        120                 125                 130

tcc ctg gga cct gga acg aag ctt cca tcg ctg ggc cat gca ccg tat   547
Ser Leu Gly Pro Gly Thr Lys Leu Pro Ser Leu Gly His Ala Pro Tyr
    135                 140                 145

gca gat ttg cgt ggg cac tac aag gaa gca gcg ctt ggc atc atc gac   595
Ala Asp Leu Arg Gly His Tyr Lys Glu Ala Ala Leu Gly Ile Ile Asp
150                 155                 160                 165

ggt ggt ggc gat gcc ttt ttg att gag act gct cag gac ttg ctt cag   643
Gly Gly Gly Asp Ala Phe Leu Ile Glu Thr Ala Gln Asp Leu Leu Gln
                170                 175                 180

gtc aag gct gcg gtt cac ggc gtt caa gat gcc atg gct gaa ctt gat   691
Val Lys Ala Ala Val His Gly Val Gln Asp Ala Met Ala Glu Leu Asp
            185                 190                 195

aca ttc ttg ccc att att tgc cac gtc acc gta gag acc acc ggc acc   739
Thr Phe Leu Pro Ile Ile Cys His Val Thr Val Glu Thr Thr Gly Thr
        200                 205                 210

atg ctc atg ggt tct gag atc ggt gcc gcg ttg aca gcg ctg cag cca   787
Met Leu Met Gly Ser Glu Ile Gly Ala Ala Leu Thr Ala Leu Gln Pro
    215                 220                 225

ctg ggt atc gac atg att ggt ctg aac tgc gcc acc ggc cca gat gag   835
Leu Gly Ile Asp Met Ile Gly Leu Asn Cys Ala Thr Gly Pro Asp Glu
230                 235                 240                 245

atg agc gag cac ctg cgt tac ctg tcc aag cac gcc gat att cct gtg   883
Met Ser Glu His Leu Arg Tyr Leu Ser Lys His Ala Asp Ile Pro Val
                250                 255                 260

tcg gtg atg cct aac gca ggt ctt cct gtc ctg ggt aaa aac ggt gca   931
Ser Val Met Pro Asn Ala Gly Leu Pro Val Leu Gly Lys Asn Gly Ala
            265                 270                 275

gaa tac cca ctt gag gct gag gat ttg gcg cag gcg ctg gct gga ttc   979
Glu Tyr Pro Leu Glu Ala Glu Asp Leu Ala Gln Ala Leu Ala Gly Phe
        280                 285                 290

gtc tcc gaa tat ggc ctg tcc atg gtg ggt ggt tgt tgt ggc acc aca  1027
Val Ser Glu Tyr Gly Leu Ser Met Val Gly Gly Cys Cys Gly Thr Thr
    295                 300                 305
```

```
cct gag cac atc cgt gcg gtc cgc gat gcg gtg gtt ggt gtt cca gag   1075
Pro Glu His Ile Arg Ala Val Arg Asp Ala Val Val Gly Val Pro Glu
310                 315                 320                 325

cag gaa acc tcc aca ctg acc aag atc cct gca ggc cct gtt gag cag   1123
Gln Glu Thr Ser Thr Leu Thr Lys Ile Pro Ala Gly Pro Val Glu Gln
                330                 335                 340

gcc tcc cgc gag gtg gag aaa gag gac tcc gtc gcg tcg ctg tac acc   1171
Ala Ser Arg Glu Val Glu Lys Glu Asp Ser Val Ala Ser Leu Tyr Thr
            345                 350                 355

tcg gtg cca ttg tcc cag gaa acc ggc att tcc atg atc ggt gag cgc   1219
Ser Val Pro Leu Ser Gln Glu Thr Gly Ile Ser Met Ile Gly Glu Arg
        360                 365                 370

acc aac tcc aac ggt tcc aag gca ttc cgt gag gca atg ctg tct ggc   1267
Thr Asn Ser Asn Gly Ser Lys Ala Phe Arg Glu Ala Met Leu Ser Gly
    375                 380                 385

gat tgg gaa aag tgt gtg gat att gcc aag cag caa acc cgc gat ggt   1315
Asp Trp Glu Lys Cys Val Asp Ile Ala Lys Gln Gln Thr Arg Asp Gly
390                 395                 400                 405

gca cac atg ctg gat ctt tgt gtg gat tac gtg gga cga gac ggc acc   1363
Ala His Met Leu Asp Leu Cys Val Asp Tyr Val Gly Arg Asp Gly Thr
                410                 415                 420

gcc gat atg gcg acc ttg gca gca ctt ctt gct acc agc tcc act ttg   1411
Ala Asp Met Ala Thr Leu Ala Ala Leu Leu Ala Thr Ser Ser Thr Leu
            425                 430                 435

cca atc atg att gac tcc acc gag cca gag gtt att cgc aca ggc ctt   1459
Pro Ile Met Ile Asp Ser Thr Glu Pro Glu Val Ile Arg Thr Gly Leu
        440                 445                 450

gag cac ttg ggt gga cga agc atc gtt aac tcc gtc aac ttt gaa gac   1507
Glu His Leu Gly Gly Arg Ser Ile Val Asn Ser Val Asn Phe Glu Asp
    455                 460                 465

ggc gat ggc cct gag tcc cgc tac cag cgc atc atg aaa ctg gta aag   1555
Gly Asp Gly Pro Glu Ser Arg Tyr Gln Arg Ile Met Lys Leu Val Lys
470                 475                 480                 485

cag cac ggt gcg gcc gtg gtt gcg ctg acc att gat gag gaa ggc cag   1603
Gln His Gly Ala Ala Val Val Ala Leu Thr Ile Asp Glu Glu Gly Gln
                490                 495                 500

gca cgt acc gct gag cac aag gtg cgc att gct aaa cga ctg att gac   1651
Ala Arg Thr Ala Glu His Lys Val Arg Ile Ala Lys Arg Leu Ile Asp
            505                 510                 515
```

```
gat atc acc ggc agc tac ggc ctg gat atc aaa gac atc gtt gtg gac   1699
Asp Ile Thr Gly Ser Tyr Gly Leu Asp Ile Lys Asp Ile Val Val Asp
        520                 525                 530

tgc ctg acc ttc ccg atc tct act ggc cag gaa gaa acc agg cga gat   1747
Cys Leu Thr Phe Pro Ile Ser Thr Gly Gln Glu Glu Thr Arg Arg Asp
    535                 540                 545

ggc att gaa acc atc gaa gcc atc cgc gag ctg aag aag ctc tac cca   1795
Gly Ile Glu Thr Ile Glu Ala Ile Arg Glu Leu Lys Lys Leu Tyr Pro
550                 555                 560                 565

gaa atc cac acc acc ctg ggt ctg tcc aat att tcc ttc ggc ctg aac   1843
Glu Ile His Thr Thr Leu Gly Leu Ser Asn Ile Ser Phe Gly Leu Asn
                570                 575                 580

cct gct gca cgc cag gtt ctt aac tct gtg ttc ctc aat gag tgc att   1891
Pro Ala Ala Arg Gln Val Leu Asn Ser Val Phe Leu Asn Glu Cys Ile
            585                 590                 595

gag gct ggt ctg gac tct gcg att gcg cac agc tcc aag att ttg ccg   1939
Glu Ala Gly Leu Asp Ser Ala Ile Ala His Ser Ser Lys Ile Leu Pro
        600                 605                 610

atg aac cgc att gat gat cgc cag cgc gaa gtg gcg ttg gat atg gtc   1987
Met Asn Arg Ile Asp Asp Arg Gln Arg Glu Val Ala Leu Asp Met Val
    615                 620                 625

tat gat cgc cgc acc gag gat tac gat ccg ctg cag gaa ttc atg cag   2035
Tyr Asp Arg Arg Thr Glu Asp Tyr Asp Pro Leu Gln Glu Phe Met Gln
630                 635                 640                 645

ctg ttt gag ggc gtt tct gct gcc gat gcc aag gat gct cgc gct gaa   2083
Leu Phe Glu Gly Val Ser Ala Ala Asp Ala Lys Asp Ala Arg Ala Glu
                650                 655                 660

cag ctg gcc gct atg cct ttg ttt gag cgt ttg gca cag cgc atc atc   2131
Gln Leu Ala Ala Met Pro Leu Phe Glu Arg Leu Ala Gln Arg Ile Ile
            665                 670                 675

gac ggc gat aag aat ggc ctt gag gat gat ctg gaa gca ggc atg aag   2179
Asp Gly Asp Lys Asn Gly Leu Glu Asp Asp Leu Glu Ala Gly Met Lys
        680                 685                 690

gag aag tct cct att gcg atc atc aac gag gac ctt ctc aac ggc atg   2227
Glu Lys Ser Pro Ile Ala Ile Ile Asn Glu Asp Leu Leu Asn Gly Met
    695                 700                 705

aag acc gtg ggt gag ctg ttt ggt tcc gga cag atg cag ctg cca ttc   2275
Lys Thr Val Gly Glu Leu Phe Gly Ser Gly Gln Met Gln Leu Pro Phe
710                 715                 720                 725
```

```
gtg ctg caa tcg gca gaa acc atg aaa act gcg gtg gcc tat ttg gaa   2323
Val Leu Gln Ser Ala Glu Thr Met Lys Thr Ala Val Ala Tyr Leu Glu
                730                 735                 740

ccg ttc atg gaa gag gaa gca gaa gct acc gga tct gcg cag gca gag   2371
Pro Phe Met Glu Glu Glu Ala Glu Ala Thr Gly Ser Ala Gln Ala Glu
            745                 750                 755

ggc aag ggc aaa atc gtc gtg gcc acc gtc aag ggt gac gtg cac gat   2419
Gly Lys Gly Lys Ile Val Val Ala Thr Val Lys Gly Asp Val His Asp
        760                 765                 770

atc ggc aag aac ttg gtg gac atc att ttg tcc aac aac ggt tac gac   2467
Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly Tyr Asp
    775                 780                 785

gtg gtg aac ttg ggc atc aag cag cca ctg tcc gcc atg ttg gaa gca   2515
Val Val Asn Leu Gly Ile Lys Gln Pro Leu Ser Ala Met Leu Glu Ala
790                 795                 800                 805

gcg gaa gaa cac aaa gca gac gtc atc ggc atg tcg gga ctt ctt gtg   2563
Ala Glu Glu His Lys Ala Asp Val Ile Gly Met Ser Gly Leu Leu Val
                810                 815                 820

aag tcc acc gtg gtg atg aag gaa aac ctt gag gag atk aac aac gcc   2611
Lys Ser Thr Val Val Met Lys Glu Asn Leu Glu Glu Xaa Asn Asn Ala
            825                 830                 835

ggc gca tcc aat tac cca gtc att ttg ggt ggc gct gcg ctg acg cgt   2659
Gly Ala Ser Asn Tyr Pro Val Ile Leu Gly Gly Ala Ala Leu Thr Arg
        840                 845                 850

acc tac gtg gaa aac gat ctc aac gag gtg tac acc ggt gag gtg tac   2707
Thr Tyr Val Glu Asn Asp Leu Asn Glu Val Tyr Thr Gly Glu Val Tyr
    855                 860                 865

tac gcc cgt gat gct ttc gag ggc ctg cgc ctg atg gat gag gtg atg   2755
Tyr Ala Arg Asp Ala Phe Glu Gly Leu Arg Leu Met Asp Glu Val Met
870                 875                 880                 885

gca gaa aag cgt ggt gaa gga ctt gat ccc aac tca cca gaa gct att   2803
Ala Glu Lys Arg Gly Glu Gly Leu Asp Pro Asn Ser Pro Glu Ala Ile
                890                 895                 900

gag cag gcg aag aag aag gcg gaa cgt aag gct cgt aat gag cgt tcc   2851
Glu Gln Ala Lys Lys Lys Ala Glu Arg Lys Ala Arg Asn Glu Arg Ser
            905                 910                 915

cgc aag att gcc gcg gag cgt aaa gct aat gcg gct ccc gtg att gtt   2899
Arg Lys Ile Ala Ala Glu Arg Lys Ala Asn Ala Ala Pro Val Ile Val
        920                 925                 930
```

```
ccg gag cgt tct gat gtc tcc acc gat act cca acc gcg gca cca ccg   2947
Pro Glu Arg Ser Asp Val Ser Thr Asp Thr Pro Thr Ala Ala Pro Pro
    935                 940                 945

ttc tgg gga acc cgc att gtc aag ggt ctg ccc ttg gcg gag ttc ttg   2995
Phe Trp Gly Thr Arg Ile Val Lys Gly Leu Pro Leu Ala Glu Phe Leu
950                 955                 960                 965

ggc aac ctt gat gag cgc gcc ttg ttc atg ggg cag tgg ggt ctg aaa   3043
Gly Asn Leu Asp Glu Arg Ala Leu Phe Met Gly Gln Trp Gly Leu Lys
                970                 975                 980

tcc acc cgc ggc aac gag ggt cca agc tat gag gat ttg gtg gaa act   3091
Ser Thr Arg Gly Asn Glu Gly Pro Ser Tyr Glu Asp Leu Val Glu Thr
            985                 990                 995

gaa ggc cga cca cgc ctg cgc tac tgg ctg gat cgc ctg aag tct gag   3139
Glu Gly Arg Pro Arg Leu Arg Tyr Trp Leu Asp Arg Leu Lys Ser Glu
       1000                1005                1010

ggc att ttg gac cac gtg gcc ttg gtg tat ggc tac ttc cca gcg gtc   3187
Gly Ile Leu Asp His Val Ala Leu Val Tyr Gly Tyr Phe Pro Ala Val
   1015                1020                1025

gcg gaa ggc gat gac gtg gtg atc ttg gaa tcc ccg gat cca cac gca   3235
Ala Glu Gly Asp Asp Val Val Ile Leu Glu Ser Pro Asp Pro His Ala
1030                1035                1040                1045

gcc gaa cgc atg cgc ttt agc ttc cca cgc cag cag cgc ggc agg ttc   3283
Ala Glu Arg Met Arg Phe Ser Phe Pro Arg Gln Gln Arg Gly Arg Phe
            1050                1055                1060

ttg tgc atc gcg gat ttc att cgc cca cgc gag caa gct gtc aag gac   3331
Leu Cys Ile Ala Asp Phe Ile Arg Pro Arg Glu Gln Ala Val Lys Asp
        1065                1070                1075

ggc caa gtg gac gtc atg cca ttc cag ctg gtc acc atg ggt aat cct   3379
Gly Gln Val Asp Val Met Pro Phe Gln Leu Val Thr Met Gly Asn Pro
       1080                1085                1090

att gct gat ttc gcc aac gag ttg ttc gca gcc aat gaa tac cgc gag   3427
Ile Ala Asp Phe Ala Asn Glu Leu Phe Ala Ala Asn Glu Tyr Arg Glu
   1095                1100                1105

tac ttg gaa gtt cac ggc atc ggc gtg cag ctc acc gaa gca ttg gcc   3475
Tyr Leu Glu Val His Gly Ile Gly Val Gln Leu Thr Glu Ala Leu Ala
1110                1115                1120                1125

gag tac tgg cac tcc cga gtg cgc agc gaa ctc aag ctg aac gac ggt   3523
Glu Tyr Trp His Ser Arg Val Arg Ser Glu Leu Lys Leu Asn Asp Gly
            1130                1135                1140
```

```
gga tct gtc gct gat ttt gat cca gaa gac aag acc aag ttc ttc gac   3571
Gly Ser Val Ala Asp Phe Asp Pro Glu Asp Lys Thr Lys Phe Phe Asp
          1145                1150                1155

ctg gat tac cgc ggc gcc cgc ttc tcc ttt ggt tac ggt tct tgc cct   3619
Leu Asp Tyr Arg Gly Ala Arg Phe Ser Phe Gly Tyr Gly Ser Cys Pro
      1160                1165                1170

gat ctg gaa gac cgc gca aag ctg gtg gaa ttg ctc gag cca ggc cgt   3667
Asp Leu Glu Asp Arg Ala Lys Leu Val Glu Leu Leu Glu Pro Gly Arg
   1175                1180                1185

atc ggc gtg gag ttg tcc gag gaa ctc cag ctg cac cca gag cag tcc   3715
Ile Gly Val Glu Leu Ser Glu Glu Leu Gln Leu His Pro Glu Gln Ser
1190                1195                1200                1205

aca gac gcg ttt gtg ctc tac cac cca gag gca aag tac ttt aac gtc   3763
Thr Asp Ala Phe Val Leu Tyr His Pro Glu Ala Lys Tyr Phe Asn Val
             1210                1215                1220

taacaccttt gagagggaaa actttcccgc                                 3793
```

<210> 2
<211> 1221
<212> PRT
<213> Corynebacterium glutamicum

<400> 2

```
Met Ser Thr Ser Val Thr Ser Pro Ala His Asn Asn Ala His Ser Ser
  1               5                  10                  15

Glu Phe Leu Asp Ala Leu Ala Asn His Val Leu Ile Gly Asp Gly Ala
             20                  25                  30

Met Gly Thr Gln Leu Gln Gly Phe Asp Leu Asp Val Glu Lys Asp Phe
         35                  40                  45

Leu Asp Leu Glu Gly Cys Asn Glu Ile Leu Asn Asp Thr Arg Pro Asp
     50                  55                  60

Val Leu Arg Gln Ile His Arg Ala Tyr Phe Glu Ala Gly Ala Asp Leu
 65                  70                  75                  80

Val Glu Thr Asn Thr Phe Gly Cys Asn Leu Pro Asn Leu Ala Asp Tyr
                 85                  90                  95

Asp Ile Ala Asp Arg Cys Arg Glu Leu Ala Tyr Lys Gly Thr Ala Val
            100                 105                 110

Ala Arg Glu Val Ala Asp Glu Met Gly Pro Gly Arg Asn Gly Met Arg
        115                 120                 125

Arg Phe Val Val Gly Ser Leu Gly Pro Gly Thr Lys Leu Pro Ser Leu
    130                 135                 140
```

```
Gly His Ala Pro Tyr Ala Asp Leu Arg Gly His Tyr Lys Glu Ala Ala
145                 150                 155                 160

Leu Gly Ile Ile Asp Gly Gly Gly Asp Ala Phe Leu Ile Glu Thr Ala
                165                 170                 175

Gln Asp Leu Leu Gln Val Lys Ala Ala Val His Gly Val Gln Asp Ala
            180                 185                 190

Met Ala Glu Leu Asp Thr Phe Leu Pro Ile Ile Cys His Val Thr Val
        195                 200                 205

Glu Thr Thr Gly Thr Met Leu Met Gly Ser Glu Ile Gly Ala Ala Leu
    210                 215                 220

Thr Ala Leu Gln Pro Leu Gly Ile Asp Met Ile Gly Leu Asn Cys Ala
225                 230                 235                 240

Thr Gly Pro Asp Glu Met Ser Glu His Leu Arg Tyr Leu Ser Lys His
                245                 250                 255

Ala Asp Ile Pro Val Ser Val Met Pro Asn Ala Gly Leu Pro Val Leu
            260                 265                 270

Gly Lys Asn Gly Ala Glu Tyr Pro Leu Glu Ala Glu Asp Leu Ala Gln
        275                 280                 285

Ala Leu Ala Gly Phe Val Ser Glu Tyr Gly Leu Ser Met Val Gly Gly
    290                 295                 300

Cys Cys Gly Thr Thr Pro Glu His Ile Arg Ala Val Arg Asp Ala Val
305                 310                 315                 320

Val Gly Val Pro Glu Gln Glu Thr Ser Thr Leu Thr Lys Ile Pro Ala
                325                 330                 335

Gly Pro Val Glu Gln Ala Ser Arg Glu Val Glu Lys Glu Asp Ser Val
            340                 345                 350

Ala Ser Leu Tyr Thr Ser Val Pro Leu Ser Gln Glu Thr Gly Ile Ser
        355                 360                 365

Met Ile Gly Glu Arg Thr Asn Ser Asn Gly Ser Lys Ala Phe Arg Glu
    370                 375                 380

Ala Met Leu Ser Gly Asp Trp Glu Lys Cys Val Asp Ile Ala Lys Gln
385                 390                 395                 400

Gln Thr Arg Asp Gly Ala His Met Leu Asp Leu Cys Val Asp Tyr Val
                405                 410                 415

Gly Arg Asp Gly Thr Ala Asp Met Ala Thr Leu Ala Ala Leu Leu Ala
            420                 425                 430

Thr Ser Ser Thr Leu Pro Ile Met Ile Asp Ser Thr Glu Pro Glu Val
        435                 440                 445
```

Ile Arg Thr Gly Leu Glu His Leu Gly Gly Arg Ser Ile Val Asn Ser
450 455 460

Val Asn Phe Glu Asp Gly Asp Gly Pro Glu Ser Arg Tyr Gln Arg Ile
465 470 475 480

Met Lys Leu Val Lys Gln His Gly Ala Ala Val Val Ala Leu Thr Ile
485 490 495

Asp Glu Glu Gly Gln Ala Arg Thr Ala Glu His Lys Val Arg Ile Ala
500 505 510

Lys Arg Leu Ile Asp Asp Ile Thr Gly Ser Tyr Gly Leu Asp Ile Lys
515 520 525

Asp Ile Val Val Asp Cys Leu Thr Phe Pro Ile Ser Thr Gly Gln Glu
530 535 540

Glu Thr Arg Arg Asp Gly Ile Glu Thr Ile Glu Ala Ile Arg Glu Leu
545 550 555 560

Lys Lys Leu Tyr Pro Glu Ile His Thr Thr Leu Gly Leu Ser Asn Ile
565 570 575

Ser Phe Gly Leu Asn Pro Ala Ala Arg Gln Val Leu Asn Ser Val Phe
580 585 590

Leu Asn Glu Cys Ile Glu Ala Gly Leu Asp Ser Ala Ile Ala His Ser
595 600 605

Ser Lys Ile Leu Pro Met Asn Arg Ile Asp Asp Arg Gln Arg Glu Val
610 615 620

Ala Leu Asp Met Val Tyr Asp Arg Arg Thr Glu Asp Tyr Asp Pro Leu
625 630 635 640

Gln Glu Phe Met Gln Leu Phe Glu Gly Val Ser Ala Ala Asp Ala Lys
645 650 655

Asp Ala Arg Ala Glu Gln Leu Ala Ala Met Pro Leu Phe Glu Arg Leu
660 665 670

Ala Gln Arg Ile Ile Asp Gly Asp Lys Asn Gly Leu Glu Asp Asp Leu
675 680 685

Glu Ala Gly Met Lys Glu Lys Ser Pro Ile Ala Ile Ile Asn Glu Asp
690 695 700

Leu Leu Asn Gly Met Lys Thr Val Gly Glu Leu Phe Gly Ser Gly Gln
705 710 715 720

Met Gln Leu Pro Phe Val Leu Gln Ser Ala Glu Thr Met Lys Thr Ala
725 730 735

Val Ala Tyr Leu Glu Pro Phe Met Glu Glu Glu Ala Glu Ala Thr Gly
740 745 750

Ser Ala Gln Ala Glu Gly Lys Gly Lys Ile Val Val Ala Thr Val Lys
755 760 765

Gly Asp Val His Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser
770 775 780

Asn Asn Gly Tyr Asp Val Val Asn Leu Gly Ile Lys Gln Pro Leu Ser
785 790 795 800

Ala Met Leu Glu Ala Ala Glu Glu His Lys Ala Asp Val Ile Gly Met
805 810 815

Ser Gly Leu Leu Val Lys Ser Thr Val Val Met Lys Glu Asn Leu Glu
820 825 830

Glu Xaa Asn Asn Ala Gly Ala Ser Asn Tyr Pro Val Ile Leu Gly Gly
835 840 845

Ala Ala Leu Thr Arg Thr Tyr Val Glu Asn Asp Leu Asn Glu Val Tyr
850 855 860

Thr Gly Glu Val Tyr Tyr Ala Arg Asp Ala Phe Glu Gly Leu Arg Leu
865 870 875 880

Met Asp Glu Val Met Ala Glu Lys Arg Gly Glu Gly Leu Asp Pro Asn
885 890 895

Ser Pro Glu Ala Ile Glu Gln Ala Lys Lys Lys Ala Glu Arg Lys Ala
900 905 910

Arg Asn Glu Arg Ser Arg Lys Ile Ala Ala Glu Arg Lys Ala Asn Ala
915 920 925

Ala Pro Val Ile Val Pro Glu Arg Ser Asp Val Ser Thr Asp Thr Pro
930 935 940

Thr Ala Ala Pro Pro Phe Trp Gly Thr Arg Ile Val Lys Gly Leu Pro
945 950 955 960

Leu Ala Glu Phe Leu Gly Asn Leu Asp Glu Arg Ala Leu Phe Met Gly
965 970 975

Gln Trp Gly Leu Lys Ser Thr Arg Gly Asn Glu Gly Pro Ser Tyr Glu
980 985 990

Asp Leu Val Glu Thr Glu Gly Arg Pro Arg Leu Arg Tyr Trp Leu Asp
995 1000 1005

Arg Leu Lys Ser Glu Gly Ile Leu Asp His Val Ala Leu Val Tyr Gly
1010 1015 1020

Tyr Phe Pro Ala Val Ala Glu Gly Asp Asp Val Val Ile Leu Glu Ser
025 1030 1035 1040

Pro Asp Pro His Ala Ala Glu Arg Met Arg Phe Ser Phe Pro Arg Gln
1045 1050 1055

Gln Arg Gly Arg Phe Leu Cys Ile Ala Asp Phe Ile Arg Pro Arg Glu
1060 1065 1070

Gln Ala Val Lys Asp Gly Gln Val Asp Val Met Pro Phe Gln Leu Val
1075 1080 1085

Thr Met Gly Asn Pro Ile Ala Asp Phe Ala Asn Glu Leu Phe Ala Ala
1090 1095 1100

Asn Glu Tyr Arg Glu Tyr Leu Glu Val His Gly Ile Gly Val Gln Leu
105 1110 1115 1120

Thr Glu Ala Leu Ala Glu Tyr Trp His Ser Arg Val Arg Ser Glu Leu
1125 1130 1135

Lys Leu Asn Asp Gly Gly Ser Val Ala Asp Phe Asp Pro Glu Asp Lys
1140 1145 1150

Thr Lys Phe Phe Asp Leu Asp Tyr Arg Gly Ala Arg Phe Ser Phe Gly
1155 1160 1165

Tyr Gly Ser Cys Pro Asp Leu Glu Asp Arg Ala Lys Leu Val Glu Leu
1170 1175 1180

Leu Glu Pro Gly Arg Ile Gly Val Glu Leu Ser Glu Glu Leu Gln Leu
185 1190 1195 1200

His Pro Glu Gln Ser Thr Asp Ala Phe Val Leu Tyr His Pro Glu Ala
1205 1210 1215

Lys Tyr Phe Asn Val
1220

<210> 3
<211> 1981
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(1951)

<400> 3
tcaatattcc gaagaaaacc gcgcagctct ctcactagtc tcaggtgagg cgaaagtggt 60

gaaagacccg ctacgcatgg tgcgcctggc tttttagaat gtg ctg caa acc tcc 115
Val Leu Gln Thr Ser
1 5

tgg cat ttc tct atc ctg gca ggc atg act gat acc tct ccg ttg aat 163
Trp His Phe Ser Ile Leu Ala Gly Met Thr Asp Thr Ser Pro Leu Asn
10 15 20

```
tct cag ccg agt gca gat cac cac cct gat cac gcg gct cgc cca gtt   211
Ser Gln Pro Ser Ala Asp His His Pro Asp His Ala Ala Arg Pro Val
            25                  30                  35

ctt gat gcc cac ggc ttg atc gtt gag cac gaa tcg gaa gag ttt cca   259
Leu Asp Ala His Gly Leu Ile Val Glu His Glu Ser Glu Glu Phe Pro
        40                  45                  50

gtc ccc gca ccc gct ccc ggt gaa cag ccc tgg gag aag aaa aac cgc   307
Val Pro Ala Pro Ala Pro Gly Glu Gln Pro Trp Glu Lys Lys Asn Arg
    55                  60                  65

gag tgg tac aaa gac gcc gtt ttc tac gaa gtg ctg gtt cgt gcc ttc   355
Glu Trp Tyr Lys Asp Ala Val Phe Tyr Glu Val Leu Val Arg Ala Phe
70                  75                  80                  85

tac gat cca gaa ggc aac gga gtc gga tcg ttg aaa ggc ctg acc gaa   403
Tyr Asp Pro Glu Gly Asn Gly Val Gly Ser Leu Lys Gly Leu Thr Glu
                90                  95                  100

aaa ctg gat tac atc cag tgg ctc ggc gtg gat tgc att tgg atc cca   451
Lys Leu Asp Tyr Ile Gln Trp Leu Gly Val Asp Cys Ile Trp Ile Pro
            105                 110                 115

ccg ttt tat gat tcc cca ctg cgc gac ggc ggt tac gat atc cgc aac   499
Pro Phe Tyr Asp Ser Pro Leu Arg Asp Gly Gly Tyr Asp Ile Arg Asn
        120                 125                 130

ttc cgt gaa atc ctg ccc gaa ttc ggc acc gtc gat gac ttc gtg gaa   547
Phe Arg Glu Ile Leu Pro Glu Phe Gly Thr Val Asp Asp Phe Val Glu
    135                 140                 145

ctc gtt gac cac gcc cac cgc cgt ggc ctg cgt gtt atc acc gac ttg   595
Leu Val Asp His Ala His Arg Arg Gly Leu Arg Val Ile Thr Asp Leu
150                 155                 160                 165

gtc atg aat cac acc tcc gac cag cac gca tgg ttc caa gaa tcc cgg   643
Val Met Asn His Thr Ser Asp Gln His Ala Trp Phe Gln Glu Ser Arg
                170                 175                 180

cgc gac cca acc ggc ccc tac gga gat ttc tat gtg tgg agc gat gat   691
Arg Asp Pro Thr Gly Pro Tyr Gly Asp Phe Tyr Val Trp Ser Asp Asp
            185                 190                 195

ccc acc ctg tac aac gaa gcc cgc atc atc ttt gta gat aca gaa gaa   739
Pro Thr Leu Tyr Asn Glu Ala Arg Ile Ile Phe Val Asp Thr Glu Glu
        200                 205                 210

tcc aac tgg acc tat gat ccg gtg cgt ggc cag tac ttc tgg cac cgc   787
Ser Asn Trp Thr Tyr Asp Pro Val Arg Gly Gln Tyr Phe Trp His Arg
    215                 220                 225
```

```
ttc ttc tcc cac caa cca gac ctc aac tac gac aac ccc gca gtc caa   835
Phe Phe Ser His Gln Pro Asp Leu Asn Tyr Asp Asn Pro Ala Val Gln
230                 235                 240                 245

gag gcc atg cta gat gtc ttg cgt ttc tgg ctg gac ctg gga ctt gat   883
Glu Ala Met Leu Asp Val Leu Arg Phe Trp Leu Asp Leu Gly Leu Asp
                250                 255                 260

ggt ttc cga cta gat gcc gtt cct tat ctt ttt gaa cgc gaa ggc acc   931
Gly Phe Arg Leu Asp Ala Val Pro Tyr Leu Phe Glu Arg Glu Gly Thr
            265                 270                 275

aac ggc gaa aac ctc aaa gaa acc cac gat ttc ctc aaa ctg tgt cgc   979
Asn Gly Glu Asn Leu Lys Glu Thr His Asp Phe Leu Lys Leu Cys Arg
        280                 285                 290

tct gtc att gag aag gaa tac ccc ggc cga atc ctg ctc gca gaa gcc   1027
Ser Val Ile Glu Lys Glu Tyr Pro Gly Arg Ile Leu Leu Ala Glu Ala
    295                 300                 305

aac caa tgg ccc caa gat gtg gtc gaa tac ttc ggt gaa aaa gac aaa   1075
Asn Gln Trp Pro Gln Asp Val Val Glu Tyr Phe Gly Glu Lys Asp Lys
310                 315                 320                 325

ggc gat gaa tgc cac atg gcc ttc cac ttc cct ttg atg ccg cgc atc   1123
Gly Asp Glu Cys His Met Ala Phe His Phe Pro Leu Met Pro Arg Ile
                330                 335                 340

ttc atg gga gtt cgc caa ggt tca cgc acc ccg atc agt gag atc ctg   1171
Phe Met Gly Val Arg Gln Gly Ser Arg Thr Pro Ile Ser Glu Ile Leu
            345                 350                 355

gcc aac acc ccg gag att ccc aag act gcc caa tgg ggt att ttc ctg   1219
Ala Asn Thr Pro Glu Ile Pro Lys Thr Ala Gln Trp Gly Ile Phe Leu
        360                 365                 370

cgt aat cat gat gag ctc acc ctt gaa atg gtc tcc gat gag gaa cgc   1267
Arg Asn His Asp Glu Leu Thr Leu Glu Met Val Ser Asp Glu Glu Arg
    375                 380                 385

agc tac atg tac tcc caa ttc gcc tcc gaa cct cgc atg cgc gcc aac   1315
Ser Tyr Met Tyr Ser Gln Phe Ala Ser Glu Pro Arg Met Arg Ala Asn
390                 395                 400                 405

gta gga atc cgc agg cgc ctt tcc cca ctg ctt gaa ggc gac cgc aac   1363
Val Gly Ile Arg Arg Arg Leu Ser Pro Leu Leu Glu Gly Asp Arg Asn
                410                 415                 420

cag ctg gaa ctc ctt cac ggt ttg ttg ctg tct cta cct ggc tca ccc   1411
Gln Leu Glu Leu Leu His Gly Leu Leu Leu Ser Leu Pro Gly Ser Pro
            425                 430                 435
```

```
gtg ttg tat tac ggt gat gaa att ggc atg ggc gac aat atc tgg ctc   1459
Val Leu Tyr Tyr Gly Asp Glu Ile Gly Met Gly Asp Asn Ile Trp Leu
        440                 445                 450

cac gac cgc gac gga gtg cgc acc ccc atg cag tgg tcc aac gac cgc   1507
His Asp Arg Asp Gly Val Arg Thr Pro Met Gln Trp Ser Asn Asp Arg
    455                 460                 465

aac ggt ggt ttc tcc aaa gct gat cct gaa cgc ctg tac ctt cca gcg   1555
Asn Gly Gly Phe Ser Lys Ala Asp Pro Glu Arg Leu Tyr Leu Pro Ala
470                 475                 480                 485

atc caa aat gat caa tac ggc tac gcc caa gta aac gtg gaa agc caa   1603
Ile Gln Asn Asp Gln Tyr Gly Tyr Ala Gln Val Asn Val Glu Ser Gln
                490                 495                 500

ctc aac cgc gaa aac tcc ctg ctg cgc tgg ctc cga aac caa atc ctt   1651
Leu Asn Arg Glu Asn Ser Leu Leu Arg Trp Leu Arg Asn Gln Ile Leu
            505                 510                 515

atc cgc aag cag tac cgc gca ttt ggt gcc gga acc tac cgt gaa gtg   1699
Ile Arg Lys Gln Tyr Arg Ala Phe Gly Ala Gly Thr Tyr Arg Glu Val
        520                 525                 530

tcc tcc acc aat gag tca gtg ttg aca ttt tta cga gaa cac aag ggc   1747
Ser Ser Thr Asn Glu Ser Val Leu Thr Phe Leu Arg Glu His Lys Gly
    535                 540                 545

caa acc att ttg tgt gtc aac aac atg agc aaa tat cct cag gca gtc   1795
Gln Thr Ile Leu Cys Val Asn Asn Met Ser Lys Tyr Pro Gln Ala Val
550                 555                 560                 565

tcg ctt gat ttg cgt gaa ttt gca gga cac acc cct cga gag atg tcg   1843
Ser Leu Asp Leu Arg Glu Phe Ala Gly His Thr Pro Arg Glu Met Ser
                570                 575                 580

ggc ggg cag ctg ttc cct acc att gct gaa cgg gag tgg att gtc act   1891
Gly Gly Gln Leu Phe Pro Thr Ile Ala Glu Arg Glu Trp Ile Val Thr
            585                 590                 595

tta gcc cct cac gga ttc ttc tgg ttt gat ctc acc gcc gat gaa aag   1939
Leu Ala Pro His Gly Phe Phe Trp Phe Asp Leu Thr Ala Asp Glu Lys
        600                 605                 610

gac gat atg gaa tgagcattgg ccaacacatc atcaccgagc                  1981
Asp Asp Met Glu
    615
```

<210> 4
<211> 617
<212> PRT
<213> Corynebacterium glutamicum

<400> 4

```
Val Leu Gln Thr Ser Trp His Phe Ser Ile Leu Ala Gly Met Thr Asp
 1               5                  10                  15

Thr Ser Pro Leu Asn Ser Gln Pro Ser Ala Asp His His Pro Asp His
            20                  25                  30

Ala Ala Arg Pro Val Leu Asp Ala His Gly Leu Ile Val Glu His Glu
        35                  40                  45

Ser Glu Glu Phe Pro Val Pro Ala Pro Ala Pro Gly Glu Gln Pro Trp
    50                  55                  60

Glu Lys Lys Asn Arg Glu Trp Tyr Lys Asp Ala Val Phe Tyr Glu Val
65                  70                  75                  80

Leu Val Arg Ala Phe Tyr Asp Pro Glu Gly Asn Gly Val Gly Ser Leu
                85                  90                  95

Lys Gly Leu Thr Glu Lys Leu Asp Tyr Ile Gln Trp Leu Gly Val Asp
            100                 105                 110

Cys Ile Trp Ile Pro Pro Phe Tyr Asp Ser Pro Leu Arg Asp Gly Gly
        115                 120                 125

Tyr Asp Ile Arg Asn Phe Arg Glu Ile Leu Pro Glu Phe Gly Thr Val
    130                 135                 140

Asp Asp Phe Val Glu Leu Val Asp His Ala His Arg Arg Gly Leu Arg
145                 150                 155                 160

Val Ile Thr Asp Leu Val Met Asn His Thr Ser Asp Gln His Ala Trp
                165                 170                 175

Phe Gln Glu Ser Arg Arg Asp Pro Thr Gly Pro Tyr Gly Asp Phe Tyr
            180                 185                 190

Val Trp Ser Asp Asp Pro Thr Leu Tyr Asn Glu Ala Arg Ile Ile Phe
        195                 200                 205

Val Asp Thr Glu Glu Ser Asn Trp Thr Tyr Asp Pro Val Arg Gly Gln
    210                 215                 220

Tyr Phe Trp His Arg Phe Phe Ser His Gln Pro Asp Leu Asn Tyr Asp
225                 230                 235                 240

Asn Pro Ala Val Gln Glu Ala Met Leu Asp Val Leu Arg Phe Trp Leu
                245                 250                 255

Asp Leu Gly Leu Asp Gly Phe Arg Leu Asp Ala Val Pro Tyr Leu Phe
            260                 265                 270

Glu Arg Glu Gly Thr Asn Gly Glu Asn Leu Lys Glu Thr His Asp Phe
        275                 280                 285
```

Leu Lys Leu Cys Arg Ser Val Ile Glu Lys Glu Tyr Pro Gly Arg Ile
290 295 300

Leu Leu Ala Glu Ala Asn Gln Trp Pro Gln Asp Val Val Glu Tyr Phe
305 310 315 320

Gly Glu Lys Asp Lys Gly Asp Glu Cys His Met Ala Phe His Phe Pro
325 330 335

Leu Met Pro Arg Ile Phe Met Gly Val Arg Gln Gly Ser Arg Thr Pro
340 345 350

Ile Ser Glu Ile Leu Ala Asn Thr Pro Glu Ile Pro Lys Thr Ala Gln
355 360 365

Trp Gly Ile Phe Leu Arg Asn His Asp Glu Leu Thr Leu Glu Met Val
370 375 380

Ser Asp Glu Glu Arg Ser Tyr Met Tyr Ser Gln Phe Ala Ser Glu Pro
385 390 395 400

Arg Met Arg Ala Asn Val Gly Ile Arg Arg Arg Leu Ser Pro Leu Leu
405 410 415

Glu Gly Asp Arg Asn Gln Leu Glu Leu Leu His Gly Leu Leu Leu Ser
420 425 430

Leu Pro Gly Ser Pro Val Leu Tyr Tyr Gly Asp Glu Ile Gly Met Gly
435 440 445

Asp Asn Ile Trp Leu His Asp Arg Asp Gly Val Arg Thr Pro Met Gln
450 455 460

Trp Ser Asn Asp Arg Asn Gly Gly Phe Ser Lys Ala Asp Pro Glu Arg
465 470 475 480

Leu Tyr Leu Pro Ala Ile Gln Asn Asp Gln Tyr Gly Tyr Ala Gln Val
485 490 495

Asn Val Glu Ser Gln Leu Asn Arg Glu Asn Ser Leu Leu Arg Trp Leu
500 505 510

Arg Asn Gln Ile Leu Ile Arg Lys Gln Tyr Arg Ala Phe Gly Ala Gly
515 520 525

Thr Tyr Arg Glu Val Ser Ser Thr Asn Glu Ser Val Leu Thr Phe Leu
530 535 540

Arg Glu His Lys Gly Gln Thr Ile Leu Cys Val Asn Asn Met Ser Lys
545 550 555 560

Tyr Pro Gln Ala Val Ser Leu Asp Leu Arg Glu Phe Ala Gly His Thr
565 570 575

Pro Arg Glu Met Ser Gly Gly Gln Leu Phe Pro Thr Ile Ala Glu Arg
580 585 590

Glu Trp Ile Val Thr Leu Ala Pro His Gly Phe Phe Trp Phe Asp Leu
595 600 605

Thr Ala Asp Glu Lys Asp Asp Met Glu
610 615

<210> 5
<211> 1840
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (363)..(1676)

<400> 5
cagaaactgt gtgcagaaat gcatgcagaa aaaggaaagt tcgggccaag atgggtgttt 60

ctgtatgccg atgatcggat ctttgacagc tgggtatgcg acaaatcacc gagagttgtt 120

aattcttaac aatggaaaag taacattgag agatgattta taccatcctg caccatttag 180

agtggggcta gtcatacccc cataacccta gctgtacgca atcgatttca aatcagttgg 240

aaaaagtcaa gaaaattacc cgagaattaa tttataccac acagtctatt gcaatagacc 300

aagctgttca gtagggtgca tgggagaaga atttcctaat aaaaactctt aaggacctcc 360

aa atg cca aag tac gac aat tcc aat gct gac cag tgg ggc ttt gaa 407
Met Pro Lys Tyr Asp Asn Ser Asn Ala Asp Gln Trp Gly Phe Glu
1 5 10 15

acc cgc tcc att cac gca ggc cag tca gta gac gca cag acc agc gca 455
Thr Arg Ser Ile His Ala Gly Gln Ser Val Asp Ala Gln Thr Ser Ala
20 25 30

cga aac ctt ccg atc tac caa tcc acc gct ttc gtg ttc gac tcc gct 503
Arg Asn Leu Pro Ile Tyr Gln Ser Thr Ala Phe Val Phe Asp Ser Ala
35 40 45

gag cac gcc aag cag cgt ttc gca ctt gag gat cta ggc cct gtt tac 551
Glu His Ala Lys Gln Arg Phe Ala Leu Glu Asp Leu Gly Pro Val Tyr
50 55 60

tcc cgc ctc acc aac cca acc gtt gag gct ttg gaa aac cgc atc gct 599
Ser Arg Leu Thr Asn Pro Thr Val Glu Ala Leu Glu Asn Arg Ile Ala
65 70 75

tcc ctc gaa ggt ggc gtc cac gct gta gcg ttc tcc tcc gga cag gcc 647
Ser Leu Glu Gly Gly Val His Ala Val Ala Phe Ser Ser Gly Gln Ala
80 85 90 95

```
gca acc acc aac gcc att ttg aac ctg gca gga gcg ggc gac cac atc   695
Ala Thr Thr Asn Ala Ile Leu Asn Leu Ala Gly Ala Gly Asp His Ile
                100                 105                 110

gtc acc tcc cca cgc ctc tac ggt ggc acc gag act cta ttc ctt atc   743
Val Thr Ser Pro Arg Leu Tyr Gly Gly Thr Glu Thr Leu Phe Leu Ile
            115                 120                 125

act ctt aac cgc ctg ggt atc gat gtt tcc ttc gtg gaa aac ccc gac   791
Thr Leu Asn Arg Leu Gly Ile Asp Val Ser Phe Val Glu Asn Pro Asp
        130                 135                 140

gac cct gag tcc tgg cag gca gcc gtt cag cca aac acc aaa gca ttc   839
Asp Pro Glu Ser Trp Gln Ala Ala Val Gln Pro Asn Thr Lys Ala Phe
    145                 150                 155

ttc ggc gag act ttc gcc aac cca cag gca gac gtc ctg gat att cct   887
Phe Gly Glu Thr Phe Ala Asn Pro Gln Ala Asp Val Leu Asp Ile Pro
160                 165                 170                 175

gcg gtg gct gaa gtt gcg cac cgc aac agc gtt cca ctg atc atc gac   935
Ala Val Ala Glu Val Ala His Arg Asn Ser Val Pro Leu Ile Ile Asp
                180                 185                 190

aac acc atc gct acc gca gcg ctc gtg cgc ccg ctc gag ctc ggc gca   983
Asn Thr Ile Ala Thr Ala Ala Leu Val Arg Pro Leu Glu Leu Gly Ala
            195                 200                 205

gac gtt gtc gtc gct tcc ctc acc aag ttc tac acc ggc aac ggc tcc   1031
Asp Val Val Val Ala Ser Leu Thr Lys Phe Tyr Thr Gly Asn Gly Ser
        210                 215                 220

gga ctg ggc ggc gtg ctt atc gac ggc gga aag ttc gat tgg act gtc   1079
Gly Leu Gly Gly Val Leu Ile Asp Gly Gly Lys Phe Asp Trp Thr Val
    225                 230                 235

gaa aag gat gga aag cca gta ttc ccc tac ttc gtc act cca gat gct   1127
Glu Lys Asp Gly Lys Pro Val Phe Pro Tyr Phe Val Thr Pro Asp Ala
240                 245                 250                 255

gct tac cac gga ttg aag tac gca gac ctt ggt gca cca gcc ttc ggc   1175
Ala Tyr His Gly Leu Lys Tyr Ala Asp Leu Gly Ala Pro Ala Phe Gly
                260                 265                 270

ctc aag gtt cgc gtt ggc ctt cta cgc gac acc ggc tcc acc ctc tcc   1223
Leu Lys Val Arg Val Gly Leu Leu Arg Asp Thr Gly Ser Thr Leu Ser
            275                 280                 285

gca ttc aac gca tgg gct gca gtc cag ggc atc gac acc ctt tcc ctg   1271
Ala Phe Asn Ala Trp Ala Ala Val Gln Gly Ile Asp Thr Leu Ser Leu
```

```
        290                 295                 300

cgc ctg gag cgc cac aac gaa aac gcc atc aag gtt gca gaa ttc ctc   1319
Arg Leu Glu Arg His Asn Glu Asn Ala Ile Lys Val Ala Glu Phe Leu
    305                 310                 315

aac aac cac gag aag gtg gaa aag gtt aac ttc gca ggc ctg aag gat   1367
Asn Asn His Glu Lys Val Glu Lys Val Asn Phe Ala Gly Leu Lys Asp
320                 325                 330                 335

tcc cct tgg tac gca acc aag gaa aag ctt ggc ctg aag tac acc ggc   1415
Ser Pro Trp Tyr Ala Thr Lys Glu Lys Leu Gly Leu Lys Tyr Thr Gly
                340                 345                 350

tcc gtt ctc acc ttc gag atc aag ggc ggc aag gat gag gct tgg gca   1463
Ser Val Leu Thr Phe Glu Ile Lys Gly Gly Lys Asp Glu Ala Trp Ala
            355                 360                 365

ttt atc gac gcc ctg aag cta cac tcc aac ctt gca aac atc ggc gat   1511
Phe Ile Asp Ala Leu Lys Leu His Ser Asn Leu Ala Asn Ile Gly Asp
        370                 375                 380

gtt cgc tcc ctc gtt gtt cac cca gca acc acc acc cat tca cag tcc   1559
Val Arg Ser Leu Val Val His Pro Ala Thr Thr Thr His Ser Gln Ser
    385                 390                 395

gac gaa gct ggc ctg gca cgc gcg ggc gtt acc cag tcc acc gtc cgc   1607
Asp Glu Ala Gly Leu Ala Arg Ala Gly Val Thr Gln Ser Thr Val Arg
400                 405                 410                 415

ctg tcc gtt ggc atc gag acc att gat gat atc atc gct gac ctc gaa   1655
Leu Ser Val Gly Ile Glu Thr Ile Asp Asp Ile Ile Ala Asp Leu Glu
                420                 425                 430

ggc ggc ttt gct gca atc tag ctttaaatag actcacccca gtgcttaaag      1706
Gly Gly Phe Ala Ala Ile
            435

cgctgggttt ttctttttca gactcgtgag aatgcaaact agactagaca gagctgtcca 1766

tatacactgg acgaagtttt agtcttgtcc acccagaaca ggcggttatt ttcatgccca 1826

ccctcgcgcc ttca                                                   1840
```

<210> 6
<211> 437
<212> PRT
<213> Corynebacterium glutamicum

<400> 6

```
Met Pro Lys Tyr Asp Asn Ser Asn Ala Asp Gln Trp Gly Phe Glu Thr
 1               5                  10                  15

Arg Ser Ile His Ala Gly Gln Ser Val Asp Ala Gln Thr Ser Ala Arg
            20                  25                  30

Asn Leu Pro Ile Tyr Gln Ser Thr Ala Phe Val Phe Asp Ser Ala Glu
        35                  40                  45

His Ala Lys Gln Arg Phe Ala Leu Glu Asp Leu Gly Pro Val Tyr Ser
    50                  55                  60

Arg Leu Thr Asn Pro Thr Val Glu Ala Leu Glu Asn Arg Ile Ala Ser
65                  70                  75                  80

Leu Glu Gly Gly Val His Ala Val Ala Phe Ser Ser Gly Gln Ala Ala
                85                  90                  95

Thr Thr Asn Ala Ile Leu Asn Leu Ala Gly Ala Gly Asp His Ile Val
            100                 105                 110

Thr Ser Pro Arg Leu Tyr Gly Gly Thr Glu Thr Leu Phe Leu Ile Thr
        115                 120                 125

Leu Asn Arg Leu Gly Ile Asp Val Ser Phe Val Glu Asn Pro Asp Asp
    130                 135                 140

Pro Glu Ser Trp Gln Ala Ala Val Gln Pro Asn Thr Lys Ala Phe Phe
145                 150                 155                 160

Gly Glu Thr Phe Ala Asn Pro Gln Ala Asp Val Leu Asp Ile Pro Ala
                165                 170                 175

Val Ala Glu Val Ala His Arg Asn Ser Val Pro Leu Ile Ile Asp Asn
            180                 185                 190

Thr Ile Ala Thr Ala Ala Leu Val Arg Pro Leu Glu Leu Gly Ala Asp
        195                 200                 205

Val Val Val Ala Ser Leu Thr Lys Phe Tyr Thr Gly Asn Gly Ser Gly
    210                 215                 220

Leu Gly Gly Val Leu Ile Asp Gly Gly Lys Phe Asp Trp Thr Val Glu
225                 230                 235                 240

Lys Asp Gly Lys Pro Val Phe Pro Tyr Phe Val Thr Pro Asp Ala Ala
                245                 250                 255

Tyr His Gly Leu Lys Tyr Ala Asp Leu Gly Ala Pro Ala Phe Gly Leu
            260                 265                 270
```

```
Lys Val Arg Val Gly Leu Leu Arg Asp Thr Gly Ser Thr Leu Ser Ala
        275                 280                 285

Phe Asn Ala Trp Ala Ala Val Gln Gly Ile Asp Thr Leu Ser Leu Arg
    290                 295                 300

Leu Glu Arg His Asn Glu Asn Ala Ile Lys Val Ala Glu Phe Leu Asn
305                 310                 315                 320

Asn His Glu Lys Val Glu Lys Val Asn Phe Ala Gly Leu Lys Asp Ser
                325                 330                 335

Pro Trp Tyr Ala Thr Lys Glu Lys Leu Gly Leu Lys Tyr Thr Gly Ser
            340                 345                 350

Val Leu Thr Phe Glu Ile Lys Gly Gly Lys Asp Glu Ala Trp Ala Phe
        355                 360                 365

Ile Asp Ala Leu Lys Leu His Ser Asn Leu Ala Asn Ile Gly Asp Val
    370                 375                 380

Arg Ser Leu Val Val His Pro Ala Thr Thr Thr His Ser Gln Ser Asp
385                 390                 395                 400

Glu Ala Gly Leu Ala Arg Ala Gly Val Thr Gln Ser Thr Val Arg Leu
                405                 410                 415

Ser Val Gly Ile Glu Thr Ile Asp Asp Ile Ile Ala Asp Leu Glu Gly
            420                 425                 430

Gly Phe Ala Ala Ile
        435
```

<210> 7
<211> 1033
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(1006)

<400> 7
gtgcggatcg ggtatccgcg ctacacttag aggtgttaga gatcatgagt ttccacgaac 60

tgtaacgcag gattcaccaa tcaatgaaag gtcgaccgac atg agc act gaa gac 115
                                              Met Ser Thr Glu Asp
                                               1   5

att gtc gtc gta gca gta gat ggc tcg gac gcc tca aaa caa gcc gtt 163

```
Ile Val Val Val Ala Val Asp Gly Ser Asp Ala Ser Lys Gln Ala Val
                10                  15                  20

cgg tgg gct gca aat acc gcc aac aaa cgt ggc att cca ctt cgc ttg   211
Arg Trp Ala Ala Asn Thr Ala Asn Lys Arg Gly Ile Pro Leu Arg Leu
            25                  30                  35

gct tcc agc tac acc atg cct cag ttc ctc tac gca gag gga atg gtt   259
Ala Ser Ser Tyr Thr Met Pro Gln Phe Leu Tyr Ala Glu Gly Met Val
        40                  45                  50

cca cca caa gag ctt ttc gat gac ctc cag gcc gaa gcc ctg gaa aag   307
Pro Pro Gln Glu Leu Phe Asp Asp Leu Gln Ala Glu Ala Leu Glu Lys
    55                  60                  65

att aac gaa gcc cgt gac atc gcc cat gag gta gcg cca gaa atc aag   355
Ile Asn Glu Ala Arg Asp Ile Ala His Glu Val Ala Pro Glu Ile Lys
70                  75                  80                  85

atc ggg cac acc atc gct gaa ggc agt ccc atc gac atg ctg ttg gaa   403
Ile Gly His Thr Ile Ala Glu Gly Ser Pro Ile Asp Met Leu Leu Glu
                90                  95                  100

atg tct ccc gat gcc aca atg atc gtc atg ggt tcc cgc gga ctc ggc   451
Met Ser Pro Asp Ala Thr Met Ile Val Met Gly Ser Arg Gly Leu Gly
            105                 110                 115

gga ctc tcc gga atg gtc atg ggc tcc gtc tcc ggt gca gtg gtc agc   499
Gly Leu Ser Gly Met Val Met Gly Ser Val Ser Gly Ala Val Val Ser
        120                 125                 130

cac gca aag tgt cca gtc gtt gtt gtc cgt gaa gac agc gca gtc aac   547
His Ala Lys Cys Pro Val Val Val Val Arg Glu Asp Ser Ala Val Asn
    135                 140                 145

gaa gac agc aag tac ggc cca gtc gtc gtc ggt gtg gat ggc tcc gaa   595
Glu Asp Ser Lys Tyr Gly Pro Val Val Val Gly Val Asp Gly Ser Glu
150                 155                 160                 165

gtc tcc caa cag gca acc gaa tac gca ttt gcg gaa gct gaa gct cgt   643
Val Ser Gln Gln Ala Thr Glu Tyr Ala Phe Ala Glu Ala Glu Ala Arg
                170                 175                 180

ggc gcc gaa ctc gtt gca gtt cac acc tgg atg gac atg cag gta cag   691
Gly Ala Glu Leu Val Ala Val His Thr Trp Met Asp Met Gln Val Gln
            185                 190                 195

gca tca ctt gca ggt ctt gca gct gct caa cag cag tgg gat gaa gtg   739
Ala Ser Leu Ala Gly Leu Ala Ala Ala Gln Gln Gln Trp Asp Glu Val
        200                 205                 210
```

```
gaa cgt cag caa acc gac atg ctg atc gaa cgc ctc gca cca ctg gtg   787
Glu Arg Gln Gln Thr Asp Met Leu Ile Glu Arg Leu Ala Pro Leu Val
    215                 220                 225

gaa aag tac cca agt gta acc gtc aag aag atc atc acc cgt gac cgc   835
Glu Lys Tyr Pro Ser Val Thr Val Lys Lys Ile Ile Thr Arg Asp Arg
230                 235                 240                 245

cca gtt cgc gca ctt gca gaa gca tct gaa aac gcg cag ctc cta gtc   883
Pro Val Arg Ala Leu Ala Glu Ala Ser Glu Asn Ala Gln Leu Leu Val
                250                 255                 260

gtt ggt tcc cat ggt cgt ggc gga ttt aag ggc atg ctc ctt ggc tcc   931
Val Gly Ser His Gly Arg Gly Gly Phe Lys Gly Met Leu Leu Gly Ser
            265                 270                 275

acc tcc cgc gca ctg ctg caa tcc gca ccg tgc cca atg atg gtg gtt   979
Thr Ser Arg Ala Leu Leu Gln Ser Ala Pro Cys Pro Met Met Val Val
        280                 285                 290

cgc cca cct gag aag att aag aag tag tttcttttaa gtttcgatgc cccggtt 1033
Arg Pro Pro Glu Lys Ile Lys Lys
    295                 300
```

<210> 8
<211> 301
<212> PRT
<213> Corynebacterium glutamicum

<400> 8

```
Met Ser Thr Glu Asp Ile Val Val Val Ala Val Asp Gly Ser Asp Ala
  1               5                  10                  15

Ser Lys Gln Ala Val Arg Trp Ala Ala Asn Thr Ala Asn Lys Arg Gly
             20                  25                  30

Ile Pro Leu Arg Leu Ala Ser Ser Tyr Thr Met Pro Gln Phe Leu Tyr
         35                  40                  45

Ala Glu Gly Met Val Pro Pro Gln Glu Leu Phe Asp Asp Leu Gln Ala
     50                  55                  60

Glu Ala Leu Glu Lys Ile Asn Glu Ala Arg Asp Ile Ala His Glu Val
 65                  70                  75                  80

Ala Pro Glu Ile Lys Ile Gly His Thr Ile Ala Glu Gly Ser Pro Ile
                 85                  90                  95

Asp Met Leu Leu Glu Met Ser Pro Asp Ala Thr Met Ile Val Met Gly
            100                 105                 110
```

```
Ser Arg Gly Leu Gly Gly Leu Ser Gly Met Val Met Gly Ser Val Ser
        115                 120                 125

Gly Ala Val Val Ser His Ala Lys Cys Pro Val Val Val Val Arg Glu
    130                 135                 140

Asp Ser Ala Val Asn Glu Asp Ser Lys Tyr Gly Pro Val Val Val Gly
145                 150                 155                 160

Val Asp Gly Ser Glu Val Ser Gln Gln Ala Thr Glu Tyr Ala Phe Ala
                165                 170                 175

Glu Ala Glu Ala Arg Gly Ala Glu Leu Val Ala Val His Thr Trp Met
            180                 185                 190

Asp Met Gln Val Gln Ala Ser Leu Ala Gly Leu Ala Ala Ala Gln Gln
        195                 200                 205

Gln Trp Asp Glu Val Glu Arg Gln Gln Thr Asp Met Leu Ile Glu Arg
    210                 215                 220

Leu Ala Pro Leu Val Glu Lys Tyr Pro Ser Val Thr Val Lys Lys Ile
225                 230                 235                 240

Ile Thr Arg Asp Arg Pro Val Arg Ala Leu Ala Glu Ala Ser Glu Asn
                245                 250                 255

Ala Gln Leu Leu Val Val Gly Ser His Gly Arg Gly Gly Phe Lys Gly
            260                 265                 270

Met Leu Leu Gly Ser Thr Ser Arg Ala Leu Leu Gln Ser Ala Pro Cys
        275                 280                 285

Pro Met Met Val Val Arg Pro Pro Glu Lys Ile Lys Lys
    290                 295                 300
```

SEQUENCE LISTING

<210> 9
<211> 1527
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(1504)
<223> RXS00315

<400> 9
ctcatggcat ctgcgccgtt cgcgttcttg ccagtgttgg ttggtttcac cgcaaccaag 60

```
cgtttcggcg gcaatgagtt cctgggcgcc gcgtattggt atg gcg atg gtg ttc   115
                                              Met Ala Met Val Phe
                                                1               5

ccg agc ttg gtg aac ggc tac gac gtg gcc gcc acc atg gct gcg ggc   163
Pro Ser Leu Val Asn Gly Tyr Asp Val Ala Ala Thr Met Ala Ala Gly
                10                  15                  20

gaa atg cca atg tgg tcc ctg ttt ggt tta gat gtt gcc caa gcc ggt   211
Glu Met Pro Met Trp Ser Leu Phe Gly Leu Asp Val Ala Gln Ala Gly
            25                  30                  35

tac cag ggc acc gtg ctt cct gtg ctg gtg gtt tct tgg att ctg gca   259
Tyr Gln Gly Thr Val Leu Pro Val Leu Val Val Ser Trp Ile Leu Ala
        40                  45                  50

acg atc gag aag ttc ctg cac aag cga ctc aag ggc act gca gac ttc   307
Thr Ile Glu Lys Phe Leu His Lys Arg Leu Lys Gly Thr Ala Asp Phe
    55                  60                  65

ctg atc act cca gtg ctg acg ttg ctg ctc acc gga ttc ctt aca ttc   355
Leu Ile Thr Pro Val Leu Thr Leu Leu Leu Thr Gly Phe Leu Thr Phe
70                  75                  80                  85

atc gcc att ggc cca gca atg cgc tgg gtg ggc gat gtg ctg gca cac   403
Ile Ala Ile Gly Pro Ala Met Arg Trp Val Gly Asp Val Leu Ala His
                90                  95                  100

ggt cta cag gga ctt tat gat ttc ggt ggt cca gtc ggc ggt ctg ctc   451
Gly Leu Gln Gly Leu Tyr Asp Phe Gly Gly Pro Val Gly Gly Leu Leu
            105                 110                 115

ttc ggt ctg gtc tac tca cca atc gtc atc act ggt ctg cac cag tcc   499
Phe Gly Leu Val Tyr Ser Pro Ile Val Ile Thr Gly Leu His Gln Ser
        120                 125                 130

ttc ccg cca att gag ctg gag ctg ttt aac cag ggt gga tcc ttc atc   547
Phe Pro Pro Ile Glu Leu Glu Leu Phe Asn Gln Gly Gly Ser Phe Ile
    135                 140                 145

ttc gca acg gca tct atg gct aat atc gcc cag ggt gcg gca tgt ttg   595
Phe Ala Thr Ala Ser Met Ala Asn Ile Ala Gln Gly Ala Ala Cys Leu
150                 155                 160                 165

gca gtg ttc ttc ctg gcg aag agt gaa aag ctc aag ggc ctt gca ggt   643
Ala Val Phe Phe Leu Ala Lys Ser Glu Lys Leu Lys Gly Leu Ala Gly
                170                 175                 180

gct tca ggt gtc tcc gct gtt ctt ggt att acg gag cct gcg atc ttc   691
Ala Ser Gly Val Ser Ala Val Leu Gly Ile Thr Glu Pro Ala Ile Phe
            185                 190                 195
```

```
ggt gtg aac ctt cgc ctg cgc tgg ccg ttc ttc atc ggt atc ggt acc   739
Gly Val Asn Leu Arg Leu Arg Trp Pro Phe Phe Ile Gly Ile Gly Thr
        200                 205                 210

gca gct atc ggt ggc gct ttg att gca ctc ttt aat atc aag gca gtt   787
Ala Ala Ile Gly Gly Ala Leu Ile Ala Leu Phe Asn Ile Lys Ala Val
    215                 220                 225

gcg ttg ggc gct gca ggt ttc ttg ggt gtt gtt tct att gat gct cca   835
Ala Leu Gly Ala Ala Gly Phe Leu Gly Val Val Ser Ile Asp Ala Pro
230                 235                 240                 245

gat atg gtc atg ttc ttg gtg tgt gca gtt gtt acc ttc ttc atc gca   883
Asp Met Val Met Phe Leu Val Cys Ala Val Val Thr Phe Phe Ile Ala
                250                 255                 260

ttc ggc gca gcg att gct tat ggc ctt tac ttg gtt cgc cgc aac ggc   931
Phe Gly Ala Ala Ile Ala Tyr Gly Leu Tyr Leu Val Arg Arg Asn Gly
            265                 270                 275

agc att gat cca gat gca acc gct gct cca gtg cct gca gga acg acc   979
Ser Ile Asp Pro Asp Ala Thr Ala Ala Pro Val Pro Ala Gly Thr Thr
        280                 285                 290

aaa gcc gaa gca gaa gca ccc gca gaa ttt tca aac gat tcc acc atc   1027
Lys Ala Glu Ala Glu Ala Pro Ala Glu Phe Ser Asn Asp Ser Thr Ile
    295                 300                 305

atc cag gca cct ttg acc ggt gaa gct att gca ctg agc agc gtc agc   1075
Ile Gln Ala Pro Leu Thr Gly Glu Ala Ile Ala Leu Ser Ser Val Ser
310                 315                 320                 325

gat gcc atg ttt gcc agc gga aag ctt ggc tcg ggc gtt gcc atc gtc   1123
Asp Ala Met Phe Ala Ser Gly Lys Leu Gly Ser Gly Val Ala Ile Val
                330                 335                 340

cca acc aag ggg cag tta gtt tct ccg gtg agt gga aag att gtg gtg   1171
Pro Thr Lys Gly Gln Leu Val Ser Pro Val Ser Gly Lys Ile Val Val
            345                 350                 355

gca ttc cca tct ggc cat gct ttc gca gtt cgc acc aag gct gag gat   1219
Ala Phe Pro Ser Gly His Ala Phe Ala Val Arg Thr Lys Ala Glu Asp
        360                 365                 370

ggt tcc aat gtg gat atc ttg atg cac att ggt ttc gac aca gta aac   1267
Gly Ser Asn Val Asp Ile Leu Met His Ile Gly Phe Asp Thr Val Asn
    375                 380                 385

ctc aac ggc acg cac ttt aac ccg ctg aag aag cag ggc gat gaa gtc   1315
Leu Asn Gly Thr His Phe Asn Pro Leu Lys Lys Gln Gly Asp Glu Val
390                 395                 400                 405
```

```
aaa gca ggg gag ctg ctg tgt gaa ttc gat att gat gcc att aag gct   1363
Lys Ala Gly Glu Leu Leu Cys Glu Phe Asp Ile Asp Ala Ile Lys Ala
                410                 415                 420

gca ggt tat gag gta acc acg ccg att gtt gtt tcg aat tac aag aaa   1411
Ala Gly Tyr Glu Val Thr Thr Pro Ile Val Val Ser Asn Tyr Lys Lys
            425                 430                 435

acc gga cct gta aac act tac ggt ttg ggc gaa att gaa gcg gga gcc   1459
Thr Gly Pro Val Asn Thr Tyr Gly Leu Gly Glu Ile Glu Ala Gly Ala
        440                 445                 450

aac ctg ctc aac gtc gca aag aaa gaa gcg gtg cca gca aca cca       1504
Asn Leu Leu Asn Val Ala Lys Lys Glu Ala Val Pro Ala Thr Pro
    455                 460                 465

taagttgaaa ccttgagtgt tcg                                       1527
```

<210> 10
<211> 468
<212> PRT
<213> Corynebacterium glutamicum

<400> 10

```
Met Ala Met Val Phe Pro Ser Leu Val Asn Gly Tyr Asp Val Ala Ala
  1               5                  10                  15

Thr Met Ala Ala Gly Glu Met Pro Met Trp Ser Leu Phe Gly Leu Asp
            20                  25                  30

Val Ala Gln Ala Gly Tyr Gln Gly Thr Val Leu Pro Val Leu Val Val
        35                  40                  45

Ser Trp Ile Leu Ala Thr Ile Glu Lys Phe Leu His Lys Arg Leu Lys
    50                  55                  60

Gly Thr Ala Asp Phe Leu Ile Thr Pro Val Leu Thr Leu Leu Leu Thr
65                  70                  75                  80

Gly Phe Leu Thr Phe Ile Ala Ile Gly Pro Ala Met Arg Trp Val Gly
                85                  90                  95

Asp Val Leu Ala His Gly Leu Gln Gly Leu Tyr Asp Phe Gly Gly Pro
            100                 105                 110

Val Gly Gly Leu Leu Phe Gly Leu Val Tyr Ser Pro Ile Val Ile Thr
        115                 120                 125

Gly Leu His Gln Ser Phe Pro Pro Ile Glu Leu Glu Leu Phe Asn Gln
    130                 135                 140

Gly Gly Ser Phe Ile Phe Ala Thr Ala Ser Met Ala Asn Ile Ala Gln
145                 150                 155                 160
```

Gly Ala Ala Cys Leu Ala Val Phe Phe Leu Ala Lys Ser Glu Lys Leu
165 170 175

Lys Gly Leu Ala Gly Ala Ser Gly Val Ser Ala Val Leu Gly Ile Thr
180 185 190

Glu Pro Ala Ile Phe Gly Val Asn Leu Arg Leu Arg Trp Pro Phe Phe
195 200 205

Ile Gly Ile Gly Thr Ala Ala Ile Gly Gly Ala Leu Ile Ala Leu Phe
210 215 220

Asn Ile Lys Ala Val Ala Leu Gly Ala Ala Gly Phe Leu Gly Val Val
225 230 235 240

Ser Ile Asp Ala Pro Asp Met Val Met Phe Leu Val Cys Ala Val Val
245 250 255

Thr Phe Phe Ile Ala Phe Gly Ala Ala Ile Ala Tyr Gly Leu Tyr Leu
260 265 270

Val Arg Arg Asn Gly Ser Ile Asp Pro Asp Ala Thr Ala Ala Pro Val
275 280 285

Pro Ala Gly Thr Thr Lys Ala Glu Ala Glu Ala Pro Ala Glu Phe Ser
290 295 300

Asn Asp Ser Thr Ile Ile Gln Ala Pro Leu Thr Gly Glu Ala Ile Ala
305 310 315 320

Leu Ser Ser Val Ser Asp Ala Met Phe Ala Ser Gly Lys Leu Gly Ser
325 330 335

Gly Val Ala Ile Val Pro Thr Lys Gly Gln Leu Val Ser Pro Val Ser
340 345 350

Gly Lys Ile Val Val Ala Phe Pro Ser Gly His Ala Phe Ala Val Arg
355 360 365

Thr Lys Ala Glu Asp Gly Ser Asn Val Asp Ile Leu Met His Ile Gly
370 375 380

Phe Asp Thr Val Asn Leu Asn Gly Thr His Phe Asn Pro Leu Lys Lys
385 390 395 400

Gln Gly Asp Glu Val Lys Ala Gly Glu Leu Leu Cys Glu Phe Asp Ile
405 410 415

Asp Ala Ile Lys Ala Ala Gly Tyr Glu Val Thr Thr Pro Ile Val Val
420 425 430

Ser Asn Tyr Lys Lys Thr Gly Pro Val Asn Thr Tyr Gly Leu Gly Glu
435 440 445

Ile Glu Ala Gly Ala Asn Leu Leu Asn Val Ala Lys Lys Glu Ala Val
450 455 460

Pro Ala Thr Pro
465

<210> 11
<211> 2187
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(2164)
<223> RXN01299

<400> 11
cgactgcggc gtctcttcct ggcactacca ttcctcgtcc tgaccaactc gccacagctg 60

gtgcaacggt cacccaagtc aaaggattga agaatcagc atg aat agc gta aat 115
Met Asn Ser Val Asn
1 5

aat tcc tcg ctt gtc cgg ctg gat gtc gat ttc ggc gac tcc acc acg 163
Asn Ser Ser Leu Val Arg Leu Asp Val Asp Phe Gly Asp Ser Thr Thr
10 15 20

gat gtc atc aac aac ctt gcc act gtt att ttc gac gct ggc cga gct 211
Asp Val Ile Asn Asn Leu Ala Thr Val Ile Phe Asp Ala Gly Arg Ala
25 30 35

tcc tcc gcc gac gcc ctt gcc aaa gac gcg ctg gat cgt gaa gca aag 259
Ser Ser Ala Asp Ala Leu Ala Lys Asp Ala Leu Asp Arg Glu Ala Lys
40 45 50

tcc ggc acc ggc gtt cct ggt caa gtt gct atc ccc cac tgc cgt tcc 307
Ser Gly Thr Gly Val Pro Gly Gln Val Ala Ile Pro His Cys Arg Ser
55 60 65

gaa gcc gta tct gtc cct acc ttg ggc ttt gct cgc ctg agc aag ggt 355
Glu Ala Val Ser Val Pro Thr Leu Gly Phe Ala Arg Leu Ser Lys Gly
70 75 80 85

gtg gac ttc agc gga cct gat ggc gat gcc aac ttg gtg ttc ctc att 403
Val Asp Phe Ser Gly Pro Asp Gly Asp Ala Asn Leu Val Phe Leu Ile
90 95 100

gca gca cct gct ggc ggc ggc aaa gag cac ctg aag atc ctg tcc aag 451
Ala Ala Pro Ala Gly Gly Gly Lys Glu His Leu Lys Ile Leu Ser Lys
105 110 115

ctt gct cgc tcc ttg gtg aag aag gat ttc atc aag gct ctg cag gaa 499
Leu Ala Arg Ser Leu Val Lys Lys Asp Phe Ile Lys Ala Leu Gln Glu
120 125 130

```
gcc acc acc gag cag gaa atc gtc gac gtt gtc gat gcc gtg ctc aac   547
Ala Thr Thr Glu Gln Glu Ile Val Asp Val Val Asp Ala Val Leu Asn
    135                 140                 145

cca gca cca aaa acc acc gag cca gct gca gct ccg gct gcg gcg gcg   595
Pro Ala Pro Lys Thr Thr Glu Pro Ala Ala Ala Pro Ala Ala Ala Ala
150                 155                 160                 165

gtt gct gag agt ggg gcg gcg tcg aca agc gtt act cgt atc gtg gca   643
Val Ala Glu Ser Gly Ala Ala Ser Thr Ser Val Thr Arg Ile Val Ala
                170                 175                 180

atc acc gca tgc cca acc ggt atc gca cac acc tac atg gct gcg gat   691
Ile Thr Ala Cys Pro Thr Gly Ile Ala His Thr Tyr Met Ala Ala Asp
            185                 190                 195

tcc ctg acg caa aac gcg gaa ggc cgc gat gat gtg gaa ctc gtt gtg   739
Ser Leu Thr Gln Asn Ala Glu Gly Arg Asp Asp Val Glu Leu Val Val
        200                 205                 210

gag act cag ggc tct tcc gct gtc acc cca gtc gat ccg aag atc atc   787
Glu Thr Gln Gly Ser Ser Ala Val Thr Pro Val Asp Pro Lys Ile Ile
    215                 220                 225

gaa gct gcc gac gcc gtc atc ttc gcc acc gac gtg gga gtt aaa gac   835
Glu Ala Ala Asp Ala Val Ile Phe Ala Thr Asp Val Gly Val Lys Asp
230                 235                 240                 245

cgc gag cgt ttc gct ggc aag cca gtc att gaa tcc ggc gtc aag cgc   883
Arg Glu Arg Phe Ala Gly Lys Pro Val Ile Glu Ser Gly Val Lys Arg
                250                 255                 260

gcg atc aat gag cca gcc aag atg atc gac gag gcc atc gca gcc tcc   931
Ala Ile Asn Glu Pro Ala Lys Met Ile Asp Glu Ala Ile Ala Ala Ser
            265                 270                 275

aag aac cca aac gcc cgc aag gtt tcc ggt tcc ggt gtc gcg gca tct   979
Lys Asn Pro Asn Ala Arg Lys Val Ser Gly Ser Gly Val Ala Ala Ser
        280                 285                 290

gct gaa acc acc ggc gag aag ctc ggc tgg ggc aag cgc atc cag cag  1027
Ala Glu Thr Thr Gly Glu Lys Leu Gly Trp Gly Lys Arg Ile Gln Gln
    295                 300                 305

gca gtc atg acc ggc gtg tcc tac atg gtt cca ttc gta gct gcc ggc  1075
Ala Val Met Thr Gly Val Ser Tyr Met Val Pro Phe Val Ala Ala Gly
310                 315                 320                 325

ggc ctc ctg ttg gct ctc ggc ttc gca ttc ggt gga tac gac atg gcg  1123
Gly Leu Leu Leu Ala Leu Gly Phe Ala Phe Gly Gly Tyr Asp Met Ala
                330                 335                 340
```

```
aac ggc tgg caa gca atc gcc acc cag ttc tct ctg acc aac ctg cca   1171
Asn Gly Trp Gln Ala Ile Ala Thr Gln Phe Ser Leu Thr Asn Leu Pro
            345                 350                 355

ggc aac acc gtc gat gtt gac ggc gtg gcc atg acc ttc gag cgt tca   1219
Gly Asn Thr Val Asp Val Asp Gly Val Ala Met Thr Phe Glu Arg Ser
        360                 365                 370

ggc ttc ctg ttg tac ttc ggc gca gtc ctg ttc gcc acc ggc caa gca   1267
Gly Phe Leu Leu Tyr Phe Gly Ala Val Leu Phe Ala Thr Gly Gln Ala
    375                 380                 385

gcc atg ggc ttc atc gtg gca gcc ctg tct ggc tac acc gca tac gca   1315
Ala Met Gly Phe Ile Val Ala Ala Leu Ser Gly Tyr Thr Ala Tyr Ala
390                 395                 400                 405

ctt gct gga cgc cca ggc atc gcg ccg ggc ttc gtc ggt ggc gcc atc   1363
Leu Ala Gly Arg Pro Gly Ile Ala Pro Gly Phe Val Gly Gly Ala Ile
                410                 415                 420

tcc gtc acc atc ggc gct ggc ttc att ggt ggt ctg gtt acc ggt atc   1411
Ser Val Thr Ile Gly Ala Gly Phe Ile Gly Gly Leu Val Thr Gly Ile
            425                 430                 435

ttg gct ggt ctc att gcc ctg tgg att ggc tcc tgg aag gtg cca cgc   1459
Leu Ala Gly Leu Ile Ala Leu Trp Ile Gly Ser Trp Lys Val Pro Arg
        440                 445                 450

gtg gtg cag tca ctg atg cct gtg gtc atc atc ccg cta ctt acc tca   1507
Val Val Gln Ser Leu Met Pro Val Val Ile Ile Pro Leu Leu Thr Ser
    455                 460                 465

gtg gtt gtt ggt ctc gtc atg tac ctc ctg ctg ggt cgc cca ctc gca   1555
Val Val Val Gly Leu Val Met Tyr Leu Leu Leu Gly Arg Pro Leu Ala
470                 475                 480                 485

tcc atc atg act ggt ttg cag gac tgg cta tcg tca atg tcc gga agc   1603
Ser Ile Met Thr Gly Leu Gln Asp Trp Leu Ser Ser Met Ser Gly Ser
                490                 495                 500

tcc gcc atc ttg ctg ggt atc atc ttg ggc ctc atg atg tgt ttc gac   1651
Ser Ala Ile Leu Leu Gly Ile Ile Leu Gly Leu Met Met Cys Phe Asp
            505                 510                 515

ctc ggc gga cca gta aac aag gca gcc tac ctc ttt ggt acc gca ggc   1699
Leu Gly Gly Pro Val Asn Lys Ala Ala Tyr Leu Phe Gly Thr Ala Gly
        520                 525                 530

ctg tct acc ggc gac caa gct tcc atg gaa atc atg gcc gcg atc atg   1747
Leu Ser Thr Gly Asp Gln Ala Ser Met Glu Ile Met Ala Ala Ile Met
    535                 540                 545
```

```
gca gct ggc atg gtc cca cca atc gcg ttg tcc att gct acc ctg ctg   1795
Ala Ala Gly Met Val Pro Pro Ile Ala Leu Ser Ile Ala Thr Leu Leu
550                 555                 560                 565

cgc aag aag ctg ttc acc cca gca gag caa gaa aac ggc aag tct tcc   1843
Arg Lys Lys Leu Phe Thr Pro Ala Glu Gln Glu Asn Gly Lys Ser Ser
                570                 575                 580

tgg ctg ctt ggc ctg gca ttc gtc tcc gaa ggt gcc atc cca ttc gcc   1891
Trp Leu Leu Gly Leu Ala Phe Val Ser Glu Gly Ala Ile Pro Phe Ala
            585                 590                 595

gca gct gac cca ttc cgt gtg atc cca gca atg atg gct ggc ggt gca   1939
Ala Ala Asp Pro Phe Arg Val Ile Pro Ala Met Met Ala Gly Gly Ala
        600                 605                 610

acc act ggt gca atc tcc atg gca ctg ggc gtc ggc tct cgg gct cca   1987
Thr Thr Gly Ala Ile Ser Met Ala Leu Gly Val Gly Ser Arg Ala Pro
    615                 620                 625

cac ggc ggt atc ttc gtg gtc tgg gca atc gaa cca tgg tgg ggc tgg   2035
His Gly Gly Ile Phe Val Val Trp Ala Ile Glu Pro Trp Trp Gly Trp
630                 635                 640                 645

ctc atc gca ctt gca gca ggc acc atc gtg tcc acc atc gtt gtc atc   2083
Leu Ile Ala Leu Ala Ala Gly Thr Ile Val Ser Thr Ile Val Val Ile
                650                 655                 660

gca ctg aag cag ttc tgg cca aac aag gcc gtc gct gca gaa gtc gcg   2131
Ala Leu Lys Gln Phe Trp Pro Asn Lys Ala Val Ala Ala Glu Val Ala
            665                 670                 675

aag caa gaa gca caa caa gca gct gta aac gca taatcggacc ttgacccgat 2184
Lys Gln Glu Ala Gln Gln Ala Ala Val Asn Ala
        680                 685

gtc                                                               2187
```

<210> 12
<211> 688
<212> PRT
<213> Corynebacterium glutamicum

<400> 12

```
Met Asn Ser Val Asn Asn Ser Ser Leu Val Arg Leu Asp Val Asp Phe
  1               5                  10                  15

Gly Asp Ser Thr Thr Asp Val Ile Asn Asn Leu Ala Thr Val Ile Phe
             20                  25                  30

Asp Ala Gly Arg Ala Ser Ser Ala Asp Ala Leu Ala Lys Asp Ala Leu
         35                  40                  45

Asp Arg Glu Ala Lys Ser Gly Thr Gly Val Pro Gly Gln Val Ala Ile
     50                  55                  60
```

Pro His Cys Arg Ser Glu Ala Val Ser Val Pro Thr Leu Gly Phe Ala
65 70 75 80

Arg Leu Ser Lys Gly Val Asp Phe Ser Gly Pro Asp Gly Asp Ala Asn
85 90 95

Leu Val Phe Leu Ile Ala Ala Pro Ala Gly Gly Gly Lys Glu His Leu
100 105 110

Lys Ile Leu Ser Lys Leu Ala Arg Ser Leu Val Lys Lys Asp Phe Ile
115 120 125

Lys Ala Leu Gln Glu Ala Thr Thr Glu Gln Glu Ile Val Asp Val Val
130 135 140

Asp Ala Val Leu Asn Pro Ala Pro Lys Thr Thr Glu Pro Ala Ala Ala
145 150 155 160

Pro Ala Ala Ala Ala Val Ala Glu Ser Gly Ala Ala Ser Thr Ser Val
165 170 175

Thr Arg Ile Val Ala Ile Thr Ala Cys Pro Thr Gly Ile Ala His Thr
180 185 190

Tyr Met Ala Ala Asp Ser Leu Thr Gln Asn Ala Glu Gly Arg Asp Asp
195 200 205

Val Glu Leu Val Val Glu Thr Gln Gly Ser Ser Ala Val Thr Pro Val
210 215 220

Asp Pro Lys Ile Ile Glu Ala Ala Asp Ala Val Ile Phe Ala Thr Asp
225 230 235 240

Val Gly Val Lys Asp Arg Glu Arg Phe Ala Gly Lys Pro Val Ile Glu
245 250 255

Ser Gly Val Lys Arg Ala Ile Asn Glu Pro Ala Lys Met Ile Asp Glu
260 265 270

Ala Ile Ala Ala Ser Lys Asn Pro Asn Ala Arg Lys Val Ser Gly Ser
275 280 285

Gly Val Ala Ala Ser Ala Glu Thr Thr Gly Glu Lys Leu Gly Trp Gly
290 295 300

Lys Arg Ile Gln Gln Ala Val Met Thr Gly Val Ser Tyr Met Val Pro
305 310 315 320

Phe Val Ala Ala Gly Gly Leu Leu Leu Ala Leu Gly Phe Ala Phe Gly
325 330 335

Gly Tyr Asp Met Ala Asn Gly Trp Gln Ala Ile Ala Thr Gln Phe Ser
340 345 350

Leu Thr Asn Leu Pro Gly Asn Thr Val Asp Val Asp Gly Val Ala Met
355 360 365

Thr Phe Glu Arg Ser Gly Phe Leu Leu Tyr Phe Gly Ala Val Leu Phe
370 375 380

Ala Thr Gly Gln Ala Ala Met Gly Phe Ile Val Ala Ala Leu Ser Gly
385 390 395 400

Tyr Thr Ala Tyr Ala Leu Ala Gly Arg Pro Gly Ile Ala Pro Gly Phe
405 410 415

Val Gly Gly Ala Ile Ser Val Thr Ile Gly Ala Gly Phe Ile Gly Gly
420 425 430

Leu Val Thr Gly Ile Leu Ala Gly Leu Ile Ala Leu Trp Ile Gly Ser
435 440 445

Trp Lys Val Pro Arg Val Val Gln Ser Leu Met Pro Val Val Ile Ile
450 455 460

Pro Leu Leu Thr Ser Val Val Val Gly Leu Val Met Tyr Leu Leu Leu
465 470 475 480

Gly Arg Pro Leu Ala Ser Ile Met Thr Gly Leu Gln Asp Trp Leu Ser
485 490 495

Ser Met Ser Gly Ser Ser Ala Ile Leu Leu Gly Ile Ile Leu Gly Leu
500 505 510

Met Met Cys Phe Asp Leu Gly Gly Pro Val Asn Lys Ala Ala Tyr Leu
515 520 525

Phe Gly Thr Ala Gly Leu Ser Thr Gly Asp Gln Ala Ser Met Glu Ile
530 535 540

Met Ala Ala Ile Met Ala Ala Gly Met Val Pro Pro Ile Ala Leu Ser
545 550 555 560

Ile Ala Thr Leu Leu Arg Lys Lys Leu Phe Thr Pro Ala Glu Gln Glu
565 570 575

Asn Gly Lys Ser Ser Trp Leu Leu Gly Leu Ala Phe Val Ser Glu Gly
580 585 590

Ala Ile Pro Phe Ala Ala Ala Asp Pro Phe Arg Val Ile Pro Ala Met
595 600 605

Met Ala Gly Gly Ala Thr Thr Gly Ala Ile Ser Met Ala Leu Gly Val
610 615 620

Gly Ser Arg Ala Pro His Gly Gly Ile Phe Val Val Trp Ala Ile Glu
625 630 635 640

Pro Trp Trp Gly Trp Leu Ile Ala Leu Ala Ala Gly Thr Ile Val Ser
645 650 655

Thr Ile Val Val Ile Ala Leu Lys Gln Phe Trp Pro Asn Lys Ala Val
660 665 670

Ala Ala Glu Val Ala Lys Gln Glu Ala Gln Gln Ala Ala Val Asn Ala
675 680 685

<210> 13
<211> 416
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (1)..(393)
<223> RXA00951

<400> 13
atc caa gca atc tta gag aag gca gca gcg ccg gcg aag cag aag gct 48
Ile Gln Ala Ile Leu Glu Lys Ala Ala Ala Pro Ala Lys Gln Lys Ala
1 5 10 15

cct gct gtg gct cct gct gta aca ccc act gac gct cct gca gcc tca 96
Pro Ala Val Ala Pro Ala Val Thr Pro Thr Asp Ala Pro Ala Ala Ser
20 25 30

gtc caa tcc aaa acc cac gac aag atc ctc acc gtc tgt ggc aac ggc 144
Val Gln Ser Lys Thr His Asp Lys Ile Leu Thr Val Cys Gly Asn Gly
35 40 45

ttg ggt acc tcc ctc ttc ctc aaa aac acc ctt gag caa gtt ttc gac 192
Leu Gly Thr Ser Leu Phe Leu Lys Asn Thr Leu Glu Gln Val Phe Asp
50 55 60

acc tgg ggt tgg ggt cca tac atg acg gtg gag gca acc gac act atc 240
Thr Trp Gly Trp Gly Pro Tyr Met Thr Val Glu Ala Thr Asp Thr Ile
65 70 75 80

tcc gcc aag ggc aaa gcc aag gaa gct gat ctc atc atg acc tct ggt 288
Ser Ala Lys Gly Lys Ala Lys Glu Ala Asp Leu Ile Met Thr Ser Gly
85 90 95

gaa atc gcc cgc acg ttg ggt gat gtt gga atc ccg gtt cac gtg atc 336
Glu Ile Ala Arg Thr Leu Gly Asp Val Gly Ile Pro Val His Val Ile
100 105 110

aat gac ttc acg agc acc gat gaa atc gat gct gcg ctt cgt gaa cgc 384
Asn Asp Phe Thr Ser Thr Asp Glu Ile Asp Ala Ala Leu Arg Glu Arg
115 120 125

tac gac atc taactacttt aaaaggacga aaa 416
Tyr Asp Ile
130

<210> 14
<211> 131
<212> PRT
<213> Corynebacterium glutamicum

<400> 14

```
Ile Gln Ala Ile Leu Glu Lys Ala Ala Ala Pro Ala Lys Gln Lys Ala
 1               5                  10                  15

Pro Ala Val Ala Pro Ala Val Thr Pro Thr Asp Ala Pro Ala Ala Ser
            20                  25                  30

Val Gln Ser Lys Thr His Asp Lys Ile Leu Thr Val Cys Gly Asn Gly
        35                  40                  45

Leu Gly Thr Ser Leu Phe Leu Lys Asn Thr Leu Glu Gln Val Phe Asp
    50                  55                  60

Thr Trp Gly Trp Gly Pro Tyr Met Thr Val Glu Ala Thr Asp Thr Ile
65                  70                  75                  80

Ser Ala Lys Gly Lys Ala Lys Glu Ala Asp Leu Ile Met Thr Ser Gly
                85                  90                  95

Glu Ile Ala Arg Thr Leu Gly Asp Val Gly Ile Pro Val His Val Ile
            100                 105                 110

Asn Asp Phe Thr Ser Thr Asp Glu Ile Asp Ala Ala Leu Arg Glu Arg
        115                 120                 125

Tyr Asp Ile
    130
```

<210> 15
<211> 1827
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(1804)
<223> RXN01244

<400> 15

```
gatatgtgtt tgtttgtcaa tatccaaatg tttgaatagt tgcacaactg ttggttttgt 60

ggtgatcttg aggaaattaa ctcaatgatt gtgaggatgg gtg gct act gtg gct   115
                                              Val Ala Thr Val Ala
                                                1               5

gat gtg aat caa gac act gta ctg aag ggc acc ggc gtt gtc ggt gga   163
Asp Val Asn Gln Asp Thr Val Leu Lys Gly Thr Gly Val Val Gly Gly
                10                  15                  20

gtc cgt tat gca agc gcg gtg tgg att acc cca cgc ccc gaa cta ccc   211
Val Arg Tyr Ala Ser Ala Val Trp Ile Thr Pro Arg Pro Glu Leu Pro
            25                  30                  35
```

```
caa gca ggc gaa gtc gtc gcc gaa gaa aac cgt gaa gca gag cag gag   259
Gln Ala Gly Glu Val Val Ala Glu Glu Asn Arg Glu Ala Glu Gln Glu
         40                  45                  50

cgt ttc gac gcc gct gca gcc aca gtc tct tct cgt ttg ctt gag cgc   307
Arg Phe Asp Ala Ala Ala Ala Thr Val Ser Ser Arg Leu Leu Glu Arg
     55                  60                  65

tcc gaa gct gct gaa gga cca gca gct gag gtg ctt aaa gct act gct   355
Ser Glu Ala Ala Glu Gly Pro Ala Ala Glu Val Leu Lys Ala Thr Ala
 70                  75                  80                  85

ggc atg gtc aat gac cgt ggc tgg cgt aag gct gtc atc aag ggt gtc   403
Gly Met Val Asn Asp Arg Gly Trp Arg Lys Ala Val Ile Lys Gly Val
                 90                  95                 100

aag ggt ggt cac cct gcg gaa tac gcc gtg gtt gca gca aca acc aag   451
Lys Gly Gly His Pro Ala Glu Tyr Ala Val Val Ala Ala Thr Thr Lys
            105                 110                 115

ttc atc tcc atg ttc gaa gcc gca ggc ggc ctg atc gcg gag cgc acc   499
Phe Ile Ser Met Phe Glu Ala Ala Gly Gly Leu Ile Ala Glu Arg Thr
        120                 125                 130

aca gac ttg cgc gac atc cgc gac cgc gtc atc gca gaa ctt cgt ggc   547
Thr Asp Leu Arg Asp Ile Arg Asp Arg Val Ile Ala Glu Leu Arg Gly
    135                 140                 145

gat gaa gag cca ggt ctg cca gct gtt tcc gga cag gtc att ctc ttt   595
Asp Glu Glu Pro Gly Leu Pro Ala Val Ser Gly Gln Val Ile Leu Phe
150                 155                 160                 165

gca gat gac ctc tcc cca gca gac acc gcg gca cta gac aca gat ctc   643
Ala Asp Asp Leu Ser Pro Ala Asp Thr Ala Ala Leu Asp Thr Asp Leu
                170                 175                 180

ttt gtg gga ctt gtc act gag ctg ggt ggc cca acg agc cac acc gcg   691
Phe Val Gly Leu Val Thr Glu Leu Gly Gly Pro Thr Ser His Thr Ala
            185                 190                 195

atc atc gca cgc cag ctc aac gtg cct tgc atc gtc gca tcc ggc gcc   739
Ile Ile Ala Arg Gln Leu Asn Val Pro Cys Ile Val Ala Ser Gly Ala
        200                 205                 210

ggc atc aag gac atc aag tcc ggc gaa aag gtg ctt atc gac ggc agc   787
Gly Ile Lys Asp Ile Lys Ser Gly Glu Lys Val Leu Ile Asp Gly Ser
    215                 220                 225

ctc ggc acc att gac cgc aac gcg gac gaa gct gaa gca acc aag ctc   835
Leu Gly Thr Ile Asp Arg Asn Ala Asp Glu Ala Glu Ala Thr Lys Leu
230                 235                 240                 245
```

```
gtc tcc gag tcc ctc gag cgc gct gct cgc atc gcc gag tgg aag ggt   883
Val Ser Glu Ser Leu Glu Arg Ala Ala Arg Ile Ala Glu Trp Lys Gly
                250                 255                 260

cct gca caa acc aag gac ggc tac cgc gtt cag ctg ttg gcc aac gtc   931
Pro Ala Gln Thr Lys Asp Gly Tyr Arg Val Gln Leu Leu Ala Asn Val
            265                 270                 275

caa gac ggc aac tct gca cag cag gct gca cag acc gaa gca gaa ggc   979
Gln Asp Gly Asn Ser Ala Gln Gln Ala Ala Gln Thr Glu Ala Glu Gly
        280                 285                 290

atc ggc ctg ttc cgc acc gaa ctg tgc ttc ctt tcc gcc acc gaa gag  1027
Ile Gly Leu Phe Arg Thr Glu Leu Cys Phe Leu Ser Ala Thr Glu Glu
    295                 300                 305

cca agc gtt gat gag cag gct gcg gtc tac tca aag gtg ctt gaa gca  1075
Pro Ser Val Asp Glu Gln Ala Ala Val Tyr Ser Lys Val Leu Glu Ala
310                 315                 320                 325

ttc cca gag tcc aag gtc gtt gtc cgc tcc ctc gac gca ggt tct gac  1123
Phe Pro Glu Ser Lys Val Val Val Arg Ser Leu Asp Ala Gly Ser Asp
                330                 335                 340

aag cca gtt cca ttc gca tcg atg gct gat gag atg aac cca gca ctg  1171
Lys Pro Val Pro Phe Ala Ser Met Ala Asp Glu Met Asn Pro Ala Leu
            345                 350                 355

ggt gtt cgt ggc ctg cgt atc gca cgt gga cag gtt gat ctg ctg act  1219
Gly Val Arg Gly Leu Arg Ile Ala Arg Gly Gln Val Asp Leu Leu Thr
        360                 365                 370

cgc cag ctc gac gca att gcg aag gcc agc gaa gaa ctc ggc cgt ggc  1267
Arg Gln Leu Asp Ala Ile Ala Lys Ala Ser Glu Glu Leu Gly Arg Gly
    375                 380                 385

gac gac gcc cca acc tgg gtt atg gct cca atg gtg gct acc gct tat  1315
Asp Asp Ala Pro Thr Trp Val Met Ala Pro Met Val Ala Thr Ala Tyr
390                 395                 400                 405

gaa gca aag tgg ttt gct gac atg tgc cgt gag cgt ggc cta atc gcc  1363
Glu Ala Lys Trp Phe Ala Asp Met Cys Arg Glu Arg Gly Leu Ile Ala
                410                 415                 420

ggc gcc atg atc gaa gtt cca gca gca tcc ctg atg gca gac aag atc  1411
Gly Ala Met Ile Glu Val Pro Ala Ala Ser Leu Met Ala Asp Lys Ile
            425                 430                 435

atg cct cac ctg gac ttt gtt tcc atc ggt acc aac gac ctg acc cag  1459
Met Pro His Leu Asp Phe Val Ser Ile Gly Thr Asn Asp Leu Thr Gln
        440                 445                 450
```

```
tac acc atg gca gcg gac cgc atg tct cct gag ctt gcc tac ctg acc   1507
Tyr Thr Met Ala Ala Asp Arg Met Ser Pro Glu Leu Ala Tyr Leu Thr
    455                 460                 465

gat cct tgg cag cca gca gtc ctg cgc ctg atc aag cac acc tgt gac   1555
Asp Pro Trp Gln Pro Ala Val Leu Arg Leu Ile Lys His Thr Cys Asp
470                 475                 480                 485

gaa ggt gct cgc ttt aac acc ccg gtc ggt gtt tgt ggt gaa gca gca   1603
Glu Gly Ala Arg Phe Asn Thr Pro Val Gly Val Cys Gly Glu Ala Ala
                490                 495                 500

gca gac cca ctg ttg gca act gtc ctc acc ggt ctt ggc gtg aac tcc   1651
Ala Asp Pro Leu Leu Ala Thr Val Leu Thr Gly Leu Gly Val Asn Ser
            505                 510                 515

ctg tcc gca gca tcc act gct ctc gca gca gtc ggt gca aag ctg tca   1699
Leu Ser Ala Ala Ser Thr Ala Leu Ala Ala Val Gly Ala Lys Leu Ser
        520                 525                 530

gag gtc acc ctg gaa acc tgt aag aag gca gca gaa gca gca ctt gac   1747
Glu Val Thr Leu Glu Thr Cys Lys Lys Ala Ala Glu Ala Ala Leu Asp
    535                 540                 545

gct gaa ggt gca act gaa gca cgc gat gct gta cgc gca gtg atc gac   1795
Ala Glu Gly Ala Thr Glu Ala Arg Asp Ala Val Arg Ala Val Ile Asp
550                 555                 560                 565

gca gca gtc taaaccactg ttgagctaaa aag                            1827
Ala Ala Val
```

<210> 16
<211> 568
<212> PRT
<213> Corynebacterium glutamicum

<400> 16

```
Val Ala Thr Val Ala Asp Val Asn Gln Asp Thr Val Leu Lys Gly Thr
  1               5                  10                  15

Gly Val Val Gly Gly Val Arg Tyr Ala Ser Ala Val Trp Ile Thr Pro
             20                  25                  30

Arg Pro Glu Leu Pro Gln Ala Gly Glu Val Val Ala Glu Glu Asn Arg
         35                  40                  45

Glu Ala Glu Gln Glu Arg Phe Asp Ala Ala Ala Ala Thr Val Ser Ser
     50                  55                  60

Arg Leu Leu Glu Arg Ser Glu Ala Ala Glu Gly Pro Ala Ala Glu Val
 65                  70                  75                  80

Leu Lys Ala Thr Ala Gly Met Val Asn Asp Arg Gly Trp Arg Lys Ala
                 85                  90                  95
```

```
Val Ile Lys Gly Val Lys Gly Gly His Pro Ala Glu Tyr Ala Val Val
            100                 105                 110

Ala Ala Thr Thr Lys Phe Ile Ser Met Phe Glu Ala Ala Gly Gly Leu
        115                 120                 125

Ile Ala Glu Arg Thr Thr Asp Leu Arg Asp Ile Arg Asp Arg Val Ile
    130                 135                 140

Ala Glu Leu Arg Gly Asp Glu Glu Pro Gly Leu Pro Ala Val Ser Gly
145                 150                 155                 160

Gln Val Ile Leu Phe Ala Asp Asp Leu Ser Pro Ala Asp Thr Ala Ala
                165                 170                 175

Leu Asp Thr Asp Leu Phe Val Gly Leu Val Thr Glu Leu Gly Gly Pro
            180                 185                 190

Thr Ser His Thr Ala Ile Ile Ala Arg Gln Leu Asn Val Pro Cys Ile
        195                 200                 205

Val Ala Ser Gly Ala Gly Ile Lys Asp Ile Lys Ser Gly Glu Lys Val
    210                 215                 220

Leu Ile Asp Gly Ser Leu Gly Thr Ile Asp Arg Asn Ala Asp Glu Ala
225                 230                 235                 240

Glu Ala Thr Lys Leu Val Ser Glu Ser Leu Glu Arg Ala Ala Arg Ile
                245                 250                 255

Ala Glu Trp Lys Gly Pro Ala Gln Thr Lys Asp Gly Tyr Arg Val Gln
            260                 265                 270

Leu Leu Ala Asn Val Gln Asp Gly Asn Ser Ala Gln Gln Ala Ala Gln
        275                 280                 285

Thr Glu Ala Glu Gly Ile Gly Leu Phe Arg Thr Glu Leu Cys Phe Leu
    290                 295                 300

Ser Ala Thr Glu Glu Pro Ser Val Asp Glu Gln Ala Ala Val Tyr Ser
305                 310                 315                 320

Lys Val Leu Glu Ala Phe Pro Glu Ser Lys Val Val Val Arg Ser Leu
                325                 330                 335

Asp Ala Gly Ser Asp Lys Pro Val Pro Phe Ala Ser Met Ala Asp Glu
            340                 345                 350

Met Asn Pro Ala Leu Gly Val Arg Gly Leu Arg Ile Ala Arg Gly Gln
        355                 360                 365

Val Asp Leu Leu Thr Arg Gln Leu Asp Ala Ile Ala Lys Ala Ser Glu
    370                 375                 380

Glu Leu Gly Arg Gly Asp Asp Ala Pro Thr Trp Val Met Ala Pro Met
385                 390                 395                 400
```

Val Ala Thr Ala Tyr Glu Ala Lys Trp Phe Ala Asp Met Cys Arg Glu
405 410 415

Arg Gly Leu Ile Ala Gly Ala Met Ile Glu Val Pro Ala Ala Ser Leu
420 425 430

Met Ala Asp Lys Ile Met Pro His Leu Asp Phe Val Ser Ile Gly Thr
435 440 445

Asn Asp Leu Thr Gln Tyr Thr Met Ala Ala Asp Arg Met Ser Pro Glu
450 455 460

Leu Ala Tyr Leu Thr Asp Pro Trp Gln Pro Ala Val Leu Arg Leu Ile
465 470 475 480

Lys His Thr Cys Asp Glu Gly Ala Arg Phe Asn Thr Pro Val Gly Val
485 490 495

Cys Gly Glu Ala Ala Ala Asp Pro Leu Leu Ala Thr Val Leu Thr Gly
500 505 510

Leu Gly Val Asn Ser Leu Ser Ala Ala Ser Thr Ala Leu Ala Ala Val
515 520 525

Gly Ala Lys Leu Ser Glu Val Thr Leu Glu Thr Cys Lys Lys Ala Ala
530 535 540

Glu Ala Ala Leu Asp Ala Glu Gly Ala Thr Glu Ala Arg Asp Ala Val
545 550 555 560

Arg Ala Val Ile Asp Ala Ala Val
565

<210> 17
<211> 390
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(367)
<223> RXA01300

<400> 17
gatcgacatt aaatcccctc ccttgggggg tttaactaac aaatcgctgc gccctaatcc 60

gttcggatta acggcgtagc aacacgaaag gacactttcc atg gct tcc aag act 115
Met Ala Ser Lys Thr
1 5

gta acc gtc ggt tcc tcc gtt ggc ctg cac gca cgt cca gca tcc atc 163
Val Thr Val Gly Ser Ser Val Gly Leu His Ala Arg Pro Ala Ser Ile
10 15 20

atc gct gaa gcg gct gct gag tac gac gac gaa atc ttg ctg acc ctg 211
Ile Ala Glu Ala Ala Ala Glu Tyr Asp Asp Glu Ile Leu Leu Thr Leu
25 30 35

gtt ggc tcc gat gat gac gaa gag acc gac gcg tcc tct tcc ctc atg 259
Val Gly Ser Asp Asp Asp Glu Glu Thr Asp Ala Ser Ser Ser Leu Met
40 45 50

atc atg gcg ctg ggc gca gag cac ggc aac gaa gtt acc gtc acc tcc 307
Ile Met Ala Leu Gly Ala Glu His Gly Asn Glu Val Thr Val Thr Ser
55 60 65

gac aac gct gaa gct gtt gag aag atc gct gcg ctt atc gca cag gac 355
Asp Asn Ala Glu Ala Val Glu Lys Ile Ala Ala Leu Ile Ala Gln Asp
70 75 80 85

ctt gac gct gag taaacaacgc tctgcttgtt aaa 390
Leu Asp Ala Glu

<210> 18
<211> 89
<212> PRT
<213> Corynebacterium glutamicum

<400> 18
Met Ala Ser Lys Thr Val Thr Val Gly Ser Ser Val Gly Leu His Ala
1 5 10 15

Arg Pro Ala Ser Ile Ile Ala Glu Ala Ala Ala Glu Tyr Asp Asp Glu
20 25 30

Ile Leu Leu Thr Leu Val Gly Ser Asp Asp Asp Glu Glu Thr Asp Ala
35 40 45

Ser Ser Ser Leu Met Ile Met Ala Leu Gly Ala Glu His Gly Asn Glu
50 55 60

Val Thr Val Thr Ser Asp Asn Ala Glu Ala Val Glu Lys Ile Ala Ala
65 70 75 80

Leu Ile Ala Gln Asp Leu Asp Ala Glu
85

<210> 19
<211> 508
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(508)
<223> RXN03002

<400> 19
ggaacttcga ggtgtcttcg tggggcgtac ggagatctag caagtgtggc tttatgtttg 60

accctatccg aatcaacatg cagtgaatta acatctactt atg ttt gta ctc aaa 115
Met Phe Val Leu Lys
1 5

gat ctg cta aag gca gaa cgc ata gaa ctc gac cgc acg gtc acc gat 163
Asp Leu Leu Lys Ala Glu Arg Ile Glu Leu Asp Arg Thr Val Thr Asp
10 15 20

tgg cgt gaa ggc atc cgc gcc gca ggt gta ctc cta gaa aag aca aac 211
Trp Arg Glu Gly Ile Arg Ala Ala Gly Val Leu Leu Glu Lys Thr Asn
25 30 35

agc att gat tcc gcc tac acc gat gcc atg atc gcc agc gtg gaa gaa 259
Ser Ile Asp Ser Ala Tyr Thr Asp Ala Met Ile Ala Ser Val Glu Glu
40 45 50

aaa ggc ccc tac att gtg gtc gct cca ggt ttc gct ttc gcg cac gcc 307
Lys Gly Pro Tyr Ile Val Val Ala Pro Gly Phe Ala Phe Ala His Ala
55 60 65

cgc ccc agc aga gca gtc cgc gag acc gct atg tcg tgg gtg cgc ctg 355
Arg Pro Ser Arg Ala Val Arg Glu Thr Ala Met Ser Trp Val Arg Leu
70 75 80 85

gcc tcc cct gtt tcc ttc ggt cac agt aag aat gat ccc ctc aat ctc 403
Ala Ser Pro Val Ser Phe Gly His Ser Lys Asn Asp Pro Leu Asn Leu
90 95 100

atc gtt gct ctc gct gcc aaa gat gcc acc gca cat acc caa gcg atg 451
Ile Val Ala Leu Ala Ala Lys Asp Ala Thr Ala His Thr Gln Ala Met
105 110 115

gcg gca ttg gct aaa gct tta gga aaa tac cga aag gat ctc gac gag 499
Ala Ala Leu Ala Lys Ala Leu Gly Lys Tyr Arg Lys Asp Leu Asp Glu
120 125 130

gca caa agt 508
Ala Gln Ser
135

<210> 20
<211> 136
<212> PRT
<213> Corynebacterium glutamicum

<400> 20
Met Phe Val Leu Lys Asp Leu Leu Lys Ala Glu Arg Ile Glu Leu Asp
1 5 10 15

Arg Thr Val Thr Asp Trp Arg Glu Gly Ile Arg Ala Ala Gly Val Leu
20 25 30

```
Leu Glu Lys Thr Asn Ser Ile Asp Ser Ala Tyr Thr Asp Ala Met Ile
         35                  40                  45

Ala Ser Val Glu Glu Lys Gly Pro Tyr Ile Val Val Ala Pro Gly Phe
     50                  55                  60

Ala Phe Ala His Ala Arg Pro Ser Arg Ala Val Arg Glu Thr Ala Met
 65                  70                  75                  80

Ser Trp Val Arg Leu Ala Ser Pro Val Ser Phe Gly His Ser Lys Asn
                 85                  90                  95

Asp Pro Leu Asn Leu Ile Val Ala Leu Ala Ala Lys Asp Ala Thr Ala
            100                 105                 110

His Thr Gln Ala Met Ala Ala Leu Ala Lys Ala Leu Gly Lys Tyr Arg
        115                 120                 125

Lys Asp Leu Asp Glu Ala Gln Ser
    130                 135
```

<210> 21
<211> 789
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (14)..(766)
<223> RXC00953

<400> 21

```
cttgcattcc ccaatg gcg cca cca acg gta ggc aac tac atc atg cag tcc 52
               Met Ala Pro Pro Thr Val Gly Asn Tyr Ile Met Gln Ser
                 1               5                   10

ttc act caa ggt ctg cag ttc ggc gtt gca gtt gcc gtg att ctc ttt   100
Phe Thr Gln Gly Leu Gln Phe Gly Val Ala Val Ala Val Ile Leu Phe
     15                  20                  25

ggt gtc cgc acc att ctt ggt gaa ctg gtc ccc gca ttc caa ggt att   148
Gly Val Arg Thr Ile Leu Gly Glu Leu Val Pro Ala Phe Gln Gly Ile
 30                  35                  40                  45

gct gcg aag gtt gtt ccc gga gct atc ccc gca ttg gat gca ccg atc   196
Ala Ala Lys Val Val Pro Gly Ala Ile Pro Ala Leu Asp Ala Pro Ile
                 50                  55                  60

gtg ttc ccc tac gcg cag aac gcc gtt ctc att ggt ttc ttg tct tcc   244
Val Phe Pro Tyr Ala Gln Asn Ala Val Leu Ile Gly Phe Leu Ser Ser
             65                  70                  75

ttc gtc ggt ggc ttg gtt ggc ctg act gtt ctt gca tcg tgg ctg aac   292
Phe Val Gly Gly Leu Val Gly Leu Thr Val Leu Ala Ser Trp Leu Asn
         80                  85                  90
```

```
cca gct ttt ggt gtc gcg ttg att ctg cct ggt ttg gtc ccc cac ttc   340
Pro Ala Phe Gly Val Ala Leu Ile Leu Pro Gly Leu Val Pro His Phe
     95                  100                 105

ttc act ggt ggc gcg gcg ggc gtt tac ggt aat gcc acg ggt ggt cgt   388
Phe Thr Gly Gly Ala Ala Gly Val Tyr Gly Asn Ala Thr Gly Gly Arg
110                 115                 120                 125

cga gga gca gta ttt ggc gcc ttt gcc aac ggt ctt ctg att acc ttc   436
Arg Gly Ala Val Phe Gly Ala Phe Ala Asn Gly Leu Leu Ile Thr Phe
                130                 135                 140

ctc cct gct ttc ctg ctt ggt gtg ctt ggt tcc ttc ggg tca gag aac   484
Leu Pro Ala Phe Leu Leu Gly Val Leu Gly Ser Phe Gly Ser Glu Asn
            145                 150                 155

acc act ttc ggt gat gcg gac ttt ggt tgg ttc gga atc gtt gtt ggt   532
Thr Thr Phe Gly Asp Ala Asp Phe Gly Trp Phe Gly Ile Val Val Gly
        160                 165                 170

tct gca gcc aag gtg gaa ggt gct ggc ggg ctc atc ttg ttg ctc atc   580
Ser Ala Ala Lys Val Glu Gly Ala Gly Gly Leu Ile Leu Leu Leu Ile
    175                 180                 185

atc gca gcg gtt ctt ctg ggt ggc gcg atg gtc ttc cag aag cgc gtc   628
Ile Ala Ala Val Leu Leu Gly Gly Ala Met Val Phe Gln Lys Arg Val
190                 195                 200                 205

gtg aat ggg cac tgg gat cca gct ccc aac cgt gag cgc gtg gag aag   676
Val Asn Gly His Trp Asp Pro Ala Pro Asn Arg Glu Arg Val Glu Lys
                210                 215                 220

gcg gaa gct gat gcc act cca acg gct ggg gct cgg acc tac cct aag   724
Ala Glu Ala Asp Ala Thr Pro Thr Ala Gly Ala Arg Thr Tyr Pro Lys
            225                 230                 235

att gct cct ccg gcg ggc gct cct acc cca ccg gct cga agc           766
Ile Ala Pro Pro Ala Gly Ala Pro Thr Pro Pro Ala Arg Ser
        240                 245                 250

taagatctcc aaaaccctga gat                                        789
```

<210> 22
<211> 251
<212> PRT
<213> Corynebacterium glutamicum

<400> 22

```
Met Ala Pro Pro Thr Val Gly Asn Tyr Ile Met Gln Ser Phe Thr Gln
  1               5                  10                  15

Gly Leu Gln Phe Gly Val Ala Val Ala Val Ile Leu Phe Gly Val Arg
             20                  25                  30
```

```
Thr Ile Leu Gly Glu Leu Val Pro Ala Phe Gln Gly Ile Ala Ala Lys
         35                  40                  45

Val Val Pro Gly Ala Ile Pro Ala Leu Asp Ala Pro Ile Val Phe Pro
     50                  55                  60

Tyr Ala Gln Asn Ala Val Leu Ile Gly Phe Leu Ser Ser Phe Val Gly
 65                  70                  75                  80

Gly Leu Val Gly Leu Thr Val Leu Ala Ser Trp Leu Asn Pro Ala Phe
                 85                  90                  95

Gly Val Ala Leu Ile Leu Pro Gly Leu Val Pro His Phe Phe Thr Gly
            100                 105                 110

Gly Ala Ala Gly Val Tyr Gly Asn Ala Thr Gly Gly Arg Arg Gly Ala
        115                 120                 125

Val Phe Gly Ala Phe Ala Asn Gly Leu Leu Ile Thr Phe Leu Pro Ala
    130                 135                 140

Phe Leu Leu Gly Val Leu Gly Ser Phe Gly Ser Glu Asn Thr Thr Phe
145                 150                 155                 160

Gly Asp Ala Asp Phe Gly Trp Phe Gly Ile Val Val Gly Ser Ala Ala
                165                 170                 175

Lys Val Glu Gly Ala Gly Gly Leu Ile Leu Leu Leu Ile Ile Ala Ala
            180                 185                 190

Val Leu Leu Gly Gly Ala Met Val Phe Gln Lys Arg Val Val Asn Gly
        195                 200                 205

His Trp Asp Pro Ala Pro Asn Arg Glu Arg Val Glu Lys Ala Glu Ala
    210                 215                 220

Asp Ala Thr Pro Thr Ala Gly Ala Arg Thr Tyr Pro Lys Ile Ala Pro
225                 230                 235                 240

Pro Ala Gly Ala Pro Thr Pro Pro Ala Arg Ser
                245                 250
```

<210> 23
<211> 553
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(553)
<223> RXC03001

<400> 23
cccggttcac gtgatcaatg acttcacgag caccgatgaa atcgatgctg cgcttcgtga 60

```
acgctacgac atctaactac tttaaaagga cgaaaatatt atg gac tgg tta acc   115
                                              Met Asp Trp Leu Thr
                                                1               5

att cct ctt ttc ctc gtt aat gaa atc ctt gcg gtt ccg gct ttc ctc   163
Ile Pro Leu Phe Leu Val Asn Glu Ile Leu Ala Val Pro Ala Phe Leu
                10                  15                  20

atc ggt atc atc acc gcc gtg gga ttg ggt gcc atg ggg cgt tcc gtc   211
Ile Gly Ile Ile Thr Ala Val Gly Leu Gly Ala Met Gly Arg Ser Val
            25                  30                  35

ggt cag gtt atc ggt gga gca atc aaa gca acg ttg ggc ttt ttg ctc   259
Gly Gln Val Ile Gly Gly Ala Ile Lys Ala Thr Leu Gly Phe Leu Leu
        40                  45                  50

att ggt gcg ggt gcc acg ttg gtc act gcc tcc ctg gag cca ctg ggt   307
Ile Gly Ala Gly Ala Thr Leu Val Thr Ala Ser Leu Glu Pro Leu Gly
    55                  60                  65

gcg atg atc atg ggt gcc aca ggc atg cgt ggt gtt gtc cca acg aat   355
Ala Met Ile Met Gly Ala Thr Gly Met Arg Gly Val Val Pro Thr Asn
70                  75                  80                  85

gaa gcc atc gcc gga atc gca cag gct gaa tac ggc gcg cag gtg gcg   403
Glu Ala Ile Ala Gly Ile Ala Gln Ala Glu Tyr Gly Ala Gln Val Ala
                90                  95                  100

tgg ctg atg att ctg ggc ttc gcc atc tct ttg gtg ttg gct cgt ttc   451
Trp Leu Met Ile Leu Gly Phe Ala Ile Ser Leu Val Leu Ala Arg Phe
            105                 110                 115

acc aac ctg cgt tat gtc ttg ctc aac gga cac cac gtg ctg ttg atg   499
Thr Asn Leu Arg Tyr Val Leu Leu Asn Gly His His Val Leu Leu Met
        120                 125                 130

tgc acc atg ctc acc atg gtc ttg gcc acc gga aga gtt gat gcg tgg   547
Cys Thr Met Leu Thr Met Val Leu Ala Thr Gly Arg Val Asp Ala Trp
    135                 140                 145

atc ttc                                                          553
Ile Phe
150
```

<210> 24
<211> 151
<212> PRT
<213> Corynebacterium glutamicum

<400> 24

```
Met Asp Trp Leu Thr Ile Pro Leu Phe Leu Val Asn Glu Ile Leu Ala
  1               5                  10                  15
```

Val Pro Ala Phe Leu Ile Gly Ile Ile Thr Ala Val Gly Leu Gly Ala
20 25 30

Met Gly Arg Ser Val Gly Gln Val Ile Gly Gly Ala Ile Lys Ala Thr
35 40 45

Leu Gly Phe Leu Leu Ile Gly Ala Gly Ala Thr Leu Val Thr Ala Ser
50 55 60

Leu Glu Pro Leu Gly Ala Met Ile Met Gly Ala Thr Gly Met Arg Gly
65 70 75 80

Val Val Pro Thr Asn Glu Ala Ile Ala Gly Ile Ala Gln Ala Glu Tyr
85 90 95

Gly Ala Gln Val Ala Trp Leu Met Ile Leu Gly Phe Ala Ile Ser Leu
100 105 110

Val Leu Ala Arg Phe Thr Asn Leu Arg Tyr Val Leu Leu Asn Gly His
115 120 125

His Val Leu Leu Met Cys Thr Met Leu Thr Met Val Leu Ala Thr Gly
130 135 140

Arg Val Asp Ala Trp Ile Phe
145 150

<210> 25
<211> 2172
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(2149)
<223> RXN01943

<400> 25
ccgattcttt ttcggcccaa ttcgtaacgg cgatcctctt aagtggacaa gaaagtctct 60

tgcccgcggg agacagaccc tacgtttaga aaggtttgac atg gcg tcc aaa ctg 115
Met Ala Ser Lys Leu
1 5

acg acg aca tcg caa cat att ctg gaa aac ctt ggt gga cca gac aat 163
Thr Thr Thr Ser Gln His Ile Leu Glu Asn Leu Gly Gly Pro Asp Asn
10 15 20

att act tcg atg act cac tgt gcg act cgc ctt cgc ttc caa gtg aag 211
Ile Thr Ser Met Thr His Cys Ala Thr Arg Leu Arg Phe Gln Val Lys
25 30 35

gat caa tcc att gtt gat caa caa gaa att gac tcc gac cca tca gtt 259
Asp Gln Ser Ile Val Asp Gln Gln Glu Ile Asp Ser Asp Pro Ser Val
40 45 50

```
ctt ggc gta gta ccc caa gga tcc acc ggt atg cag gtg gtg atg ggt   307
Leu Gly Val Val Pro Gln Gly Ser Thr Gly Met Gln Val Val Met Gly
     55                  60                  65

gga tct gtt gca aac tat tac caa gaa atc ctc aaa ctt gat gga atg   355
Gly Ser Val Ala Asn Tyr Tyr Gln Glu Ile Leu Lys Leu Asp Gly Met
 70                  75                  80                  85

aag cac ttc gcc gac ggt gaa gct aca gag agt tca tcc aag aag gaa   403
Lys His Phe Ala Asp Gly Glu Ala Thr Glu Ser Ser Ser Lys Lys Glu
                 90                  95                 100

tac ggc gga gtc cgt ggc aag tac tcg tgg att gac tac gcc ttc gag   451
Tyr Gly Gly Val Arg Gly Lys Tyr Ser Trp Ile Asp Tyr Ala Phe Glu
            105                 110                 115

ttc ttg tct gat act ttc cga cca atc ctg tgg gcc ctg ctt ggt gcc   499
Phe Leu Ser Asp Thr Phe Arg Pro Ile Leu Trp Ala Leu Leu Gly Ala
        120                 125                 130

tca ctg att att acc ttg ttg gtt ctt gcg gat act ttc ggt ttg caa   547
Ser Leu Ile Ile Thr Leu Leu Val Leu Ala Asp Thr Phe Gly Leu Gln
    135                 140                 145

gac ttc cgc gct cca atg gat gag cag cct gat act tat gta ttc ctg   595
Asp Phe Arg Ala Pro Met Asp Glu Gln Pro Asp Thr Tyr Val Phe Leu
150                 155                 160                 165

cac tcc atg tgg cgc tcg gtc ttc tac ttc ctg cca att atg gtt ggt   643
His Ser Met Trp Arg Ser Val Phe Tyr Phe Leu Pro Ile Met Val Gly
                170                 175                 180

gcc acc gca gct cga aag ctc ggc gca aac gag tgg att ggt gca gct   691
Ala Thr Ala Ala Arg Lys Leu Gly Ala Asn Glu Trp Ile Gly Ala Ala
            185                 190                 195

att cca gcc gca ctt ctt act cca gaa ttc ttg gca ctg ggt tct gcc   739
Ile Pro Ala Ala Leu Leu Thr Pro Glu Phe Leu Ala Leu Gly Ser Ala
        200                 205                 210

ggc gat acc gtc aca gtc ttt ggc ctg cca atg gtt ctg aat gac tac   787
Gly Asp Thr Val Thr Val Phe Gly Leu Pro Met Val Leu Asn Asp Tyr
    215                 220                 225

tcc gga cag gta ttc cca ccg ctg att gca gca att ggt ctg tac tgg   835
Ser Gly Gln Val Phe Pro Pro Leu Ile Ala Ala Ile Gly Leu Tyr Trp
230                 235                 240                 245

gtg gaa aag gga ctg aag aag atc atc cct gaa gca gtc caa atg gtg   883
Val Glu Lys Gly Leu Lys Lys Ile Ile Pro Glu Ala Val Gln Met Val
                250                 255                 260
```

```
ttc gtc cca ttc ttc tcc ctg ctg att atg atc cca gcg acc gca ttc   931
Phe Val Pro Phe Phe Ser Leu Leu Ile Met Ile Pro Ala Thr Ala Phe
            265                 270                 275

ctg ctt gga cct ttc ggc atc ggt gtt ggt aac gga att tcc aac ctg   979
Leu Leu Gly Pro Phe Gly Ile Gly Val Gly Asn Gly Ile Ser Asn Leu
        280                 285                 290

ctt gaa gcg att aac aac ttc agc cca ttt att ctt tcc atc gtt atc   1027
Leu Glu Ala Ile Asn Asn Phe Ser Pro Phe Ile Leu Ser Ile Val Ile
    295                 300                 305

cca ttg ctc tac cca ttc ttg gtt cca ctt gga ttg cac tgg cca cta   1075
Pro Leu Leu Tyr Pro Phe Leu Val Pro Leu Gly Leu His Trp Pro Leu
310                 315                 320                 325

aac gcc atc atg atc cag aac atc aac acc ctg ggt tac gac ttc att   1123
Asn Ala Ile Met Ile Gln Asn Ile Asn Thr Leu Gly Tyr Asp Phe Ile
                330                 335                 340

cag gga cca atg ggt gcc tgg aac ttc gcc tgc ttc ggc ctg gtc acc   1171
Gln Gly Pro Met Gly Ala Trp Asn Phe Ala Cys Phe Gly Leu Val Thr
            345                 350                 355

ggc gtg ttc ttg ctc tcc att aag gaa cga aac aag gcc atg cgt cag   1219
Gly Val Phe Leu Leu Ser Ile Lys Glu Arg Asn Lys Ala Met Arg Gln
        360                 365                 370

gtt tcc ctg ggt ggc atg ttg gct ggt ttg ctc ggc ggc att tcc gag   1267
Val Ser Leu Gly Gly Met Leu Ala Gly Leu Leu Gly Gly Ile Ser Glu
    375                 380                 385

cct tcc ctc tac ggt gtt ctg ctc cga ttc aag aag acc tac ttc cgc   1315
Pro Ser Leu Tyr Gly Val Leu Leu Arg Phe Lys Lys Thr Tyr Phe Arg
390                 395                 400                 405

ctc ctg ccg ggt tgt ttg gca ggc ggt atc gtg atg ggc atc ttc gac   1363
Leu Leu Pro Gly Cys Leu Ala Gly Gly Ile Val Met Gly Ile Phe Asp
                410                 415                 420

atc aag gcg tac gct ttc gtg ttc acc tcc ttg ctt acc atc cca gca   1411
Ile Lys Ala Tyr Ala Phe Val Phe Thr Ser Leu Leu Thr Ile Pro Ala
            425                 430                 435

atg gac cca tgg ttg ggc tac acc att ggt atc gca gtt gca ttc ttc   1459
Met Asp Pro Trp Leu Gly Tyr Thr Ile Gly Ile Ala Val Ala Phe Phe
        440                 445                 450

gtt tcc atg ttc ctt gtt ctc gca ctg gac tac cgt tcc aac gaa gag   1507
Val Ser Met Phe Leu Val Leu Ala Leu Asp Tyr Arg Ser Asn Glu Glu
    455                 460                 465
```

```
cgc gat gag gca cgt gca aag gtt gct gct gac aag cag gca gaa gaa   1555
Arg Asp Glu Ala Arg Ala Lys Val Ala Ala Asp Lys Gln Ala Glu Glu
470                 475                 480                 485

gat ctg aag gca gaa gct aat gca act cct gca gct cca gta gct gct   1603
Asp Leu Lys Ala Glu Ala Asn Ala Thr Pro Ala Ala Pro Val Ala Ala
                490                 495                 500

gca ggt gcg gga gcc ggt gca ggt gca gga gcc gct gct ggc gct gca   1651
Ala Gly Ala Gly Ala Gly Ala Gly Ala Gly Ala Ala Ala Gly Ala Ala
            505                 510                 515

acc gcc gtg gca gct aag ccg aag ctg gcc gct ggg gaa gta gtg gac   1699
Thr Ala Val Ala Ala Lys Pro Lys Leu Ala Ala Gly Glu Val Val Asp
        520                 525                 530

att gtt tcc cca ctc gaa ggc aag gca att cca ctt tct gaa gta cct   1747
Ile Val Ser Pro Leu Glu Gly Lys Ala Ile Pro Leu Ser Glu Val Pro
    535                 540                 545

gac cca atc ttt gca gca ggc aag ctt gga cca ggc att gca atc caa   1795
Asp Pro Ile Phe Ala Ala Gly Lys Leu Gly Pro Gly Ile Ala Ile Gln
550                 555                 560                 565

cca act gga aac acc gtt gtt gct cca gca gac gct act gtc atc ctt   1843
Pro Thr Gly Asn Thr Val Val Ala Pro Ala Asp Ala Thr Val Ile Leu
                570                 575                 580

gtc cag aaa tct gga cac gca gtg gca ttg cgc tta gat agc gga gtt   1891
Val Gln Lys Ser Gly His Ala Val Ala Leu Arg Leu Asp Ser Gly Val
            585                 590                 595

gaa atc ctt gtc cac gtt gga ttg gac acc gtg caa ttg ggc ggc gaa   1939
Glu Ile Leu Val His Val Gly Leu Asp Thr Val Gln Leu Gly Gly Glu
        600                 605                 610

ggc ttc acc gtt cac gtt gag cgc agg cag caa gtc aag gcg ggg gat   1987
Gly Phe Thr Val His Val Glu Arg Arg Gln Gln Val Lys Ala Gly Asp
    615                 620                 625

cca ctg atc act ttt gac gct gac ttc att cga tcc aag gat cta cct   2035
Pro Leu Ile Thr Phe Asp Ala Asp Phe Ile Arg Ser Lys Asp Leu Pro
630                 635                 640                 645

ttg atc acc cca gtt gtg gtg tct aac gcc gcg aaa ttc ggt gaa att   2083
Leu Ile Thr Pro Val Val Val Ser Asn Ala Ala Lys Phe Gly Glu Ile
                650                 655                 660

gaa ggt att cct gca gat cag gca aat tct tcc acg act gtg atc aag   2131
Glu Gly Ile Pro Ala Asp Gln Ala Asn Ser Ser Thr Thr Val Ile Lys
            665                 670                 675
```

gtc aac ggc aag aac gag taacctggga tccatgttgc gca 2172
Val Asn Gly Lys Asn Glu
680

<210> 26
<211> 683
<212> PRT
<213> Corynebacterium glutamicum

<400> 26
Met Ala Ser Lys Leu Thr Thr Thr Ser Gln His Ile Leu Glu Asn Leu
1 5 10 15

Gly Gly Pro Asp Asn Ile Thr Ser Met Thr His Cys Ala Thr Arg Leu
20 25 30

Arg Phe Gln Val Lys Asp Gln Ser Ile Val Asp Gln Gln Glu Ile Asp
35 40 45

Ser Asp Pro Ser Val Leu Gly Val Val Pro Gln Gly Ser Thr Gly Met
50 55 60

Gln Val Val Met Gly Gly Ser Val Ala Asn Tyr Tyr Gln Glu Ile Leu
65 70 75 80

Lys Leu Asp Gly Met Lys His Phe Ala Asp Gly Glu Ala Thr Glu Ser
85 90 95

Ser Ser Lys Lys Glu Tyr Gly Gly Val Arg Gly Lys Tyr Ser Trp Ile
100 105 110

Asp Tyr Ala Phe Glu Phe Leu Ser Asp Thr Phe Arg Pro Ile Leu Trp
115 120 125

Ala Leu Leu Gly Ala Ser Leu Ile Ile Thr Leu Leu Val Leu Ala Asp
130 135 140

Thr Phe Gly Leu Gln Asp Phe Arg Ala Pro Met Asp Glu Gln Pro Asp
145 150 155 160

Thr Tyr Val Phe Leu His Ser Met Trp Arg Ser Val Phe Tyr Phe Leu
165 170 175

Pro Ile Met Val Gly Ala Thr Ala Ala Arg Lys Leu Gly Ala Asn Glu
180 185 190

Trp Ile Gly Ala Ala Ile Pro Ala Ala Leu Leu Thr Pro Glu Phe Leu
195 200 205

Ala Leu Gly Ser Ala Gly Asp Thr Val Thr Val Phe Gly Leu Pro Met
210 215 220

Val Leu Asn Asp Tyr Ser Gly Gln Val Phe Pro Pro Leu Ile Ala Ala
225 230 235 240

```
Ile Gly Leu Tyr Trp Val Glu Lys Gly Leu Lys Lys Ile Ile Pro Glu
                245                 250                 255

Ala Val Gln Met Val Phe Val Pro Phe Phe Ser Leu Leu Ile Met Ile
            260                 265                 270

Pro Ala Thr Ala Phe Leu Leu Gly Pro Phe Gly Ile Gly Val Gly Asn
        275                 280                 285

Gly Ile Ser Asn Leu Leu Glu Ala Ile Asn Asn Phe Ser Pro Phe Ile
    290                 295                 300

Leu Ser Ile Val Ile Pro Leu Leu Tyr Pro Phe Leu Val Pro Leu Gly
305                 310                 315                 320

Leu His Trp Pro Leu Asn Ala Ile Met Ile Gln Asn Ile Asn Thr Leu
                325                 330                 335

Gly Tyr Asp Phe Ile Gln Gly Pro Met Gly Ala Trp Asn Phe Ala Cys
            340                 345                 350

Phe Gly Leu Val Thr Gly Val Phe Leu Leu Ser Ile Lys Glu Arg Asn
        355                 360                 365

Lys Ala Met Arg Gln Val Ser Leu Gly Gly Met Leu Ala Gly Leu Leu
    370                 375                 380

Gly Gly Ile Ser Glu Pro Ser Leu Tyr Gly Val Leu Leu Arg Phe Lys
385                 390                 395                 400

Lys Thr Tyr Phe Arg Leu Leu Pro Gly Cys Leu Ala Gly Gly Ile Val
                405                 410                 415

Met Gly Ile Phe Asp Ile Lys Ala Tyr Ala Phe Val Phe Thr Ser Leu
            420                 425                 430

Leu Thr Ile Pro Ala Met Asp Pro Trp Leu Gly Tyr Thr Ile Gly Ile
        435                 440                 445

Ala Val Ala Phe Phe Val Ser Met Phe Leu Val Leu Ala Leu Asp Tyr
    450                 455                 460

Arg Ser Asn Glu Glu Arg Asp Glu Ala Arg Ala Lys Val Ala Ala Asp
465                 470                 475                 480

Lys Gln Ala Glu Glu Asp Leu Lys Ala Glu Ala Asn Ala Thr Pro Ala
                485                 490                 495

Ala Pro Val Ala Ala Ala Gly Ala Gly Ala Gly Ala Gly Ala Gly Ala
            500                 505                 510

Ala Ala Gly Ala Ala Thr Ala Val Ala Ala Lys Pro Lys Leu Ala Ala
        515                 520                 525

Gly Glu Val Val Asp Ile Val Ser Pro Leu Glu Gly Lys Ala Ile Pro
    530                 535                 540
```

```
Leu Ser Glu Val Pro Asp Pro Ile Phe Ala Ala Gly Lys Leu Gly Pro
545                 550                 555                 560

Gly Ile Ala Ile Gln Pro Thr Gly Asn Thr Val Val Ala Pro Ala Asp
                565                 570                 575

Ala Thr Val Ile Leu Val Gln Lys Ser Gly His Ala Val Ala Leu Arg
            580                 585                 590

Leu Asp Ser Gly Val Glu Ile Leu Val His Val Gly Leu Asp Thr Val
        595                 600                 605

Gln Leu Gly Gly Glu Gly Phe Thr Val His Val Glu Arg Arg Gln Gln
    610                 615                 620

Val Lys Ala Gly Asp Pro Leu Ile Thr Phe Asp Ala Asp Phe Ile Arg
625                 630                 635                 640

Ser Lys Asp Leu Pro Leu Ile Thr Pro Val Val Val Ser Asn Ala Ala
                645                 650                 655

Lys Phe Gly Glu Ile Glu Gly Ile Pro Ala Asp Gln Ala Asn Ser Ser
            660                 665                 670

Thr Thr Val Ile Lys Val Asn Gly Lys Asn Glu
        675                 680
```

<210> 27
<211> 372
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(349)
<223> RXA01503

<400> 27
gtatcctcaa aggccttcta gctgttgcag ctgcagcgca ctcggtggat acgacatcca 60

cgacctatca aattctttat gctgcaggcg atgccttttc atg ttc ttg gca gtc 115
                                        Met Phe Leu Ala Val
                                          1   5

att ttg gcg att act gcg gct cgt aaa ttc ggt gcc aat gtc ttt aca 163
Ile Leu Ala Ile Thr Ala Ala Arg Lys Phe Gly Ala Asn Val Phe Thr
                10   15   20

tca gtc gca ctc gct ggt gca ttg ctg cac aca cag ctt cag gca gta 211
Ser Val Ala Leu Ala Gly Ala Leu Leu His Thr Gln Leu Gln Ala Val
            25   30   35

acc gtg ttg gtt gac ggt gaa ctc cag tcg atg act ctg gtg gct ttc 259
Thr Val Leu Val Asp Gly Glu Leu Gln Ser Met Thr Leu Val Ala Phe
        40   45   50
```
```

caa aag gct ggt aat gac gtc acc ttc ctg ggc att cca gtg gtg ctg 307
Gln Lys Ala Gly Asn Asp Val Thr Phe Leu Gly Ile Pro Val Val Leu
55 60 65

cag ttg gcg ttg cat gta gcg agt ttg atg aag ttg tcg cga 349
Gln Leu Ala Leu His Val Ala Ser Leu Met Lys Leu Ser Arg
70 75 80

taagaggagg ggcgtgtcgg tct 372

<210> 28
<211> 83
<212> PRT
<213> Corynebacterium glutamicum

<400> 28
Met Phe Leu Ala Val Ile Leu Ala Ile Thr Ala Ala Arg Lys Phe Gly
1 5 10 15

Ala Asn Val Phe Thr Ser Val Ala Leu Ala Gly Ala Leu Leu His Thr
20 25 30

Gln Leu Gln Ala Val Thr Val Leu Val Asp Gly Glu Leu Gln Ser Met
35 40 45

Thr Leu Val Ala Phe Gln Lys Ala Gly Asn Asp Val Thr Phe Leu Gly
50 55 60

Ile Pro Val Val Leu Gln Leu Ala Leu His Val Ala Ser Leu Met Lys
65 70 75 80

Leu Ser Arg

<210> 29
<211> 1578
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(1555)
<223> RXN00351

<400> 29
gaaggctgct gctaagaaaa cgaccaagaa gaccactaag aaaactacta aaaagaccac 60

cgcaaagaag accacaaaga agtcttaagc cggatcttat atg gat gat tcc aat 115
Met Asp Asp Ser Asn
1 5

agc ttt gta gtt gtt gct aac cgt ctg cca gtg gat atg act gtc cac 163
Ser Phe Val Val Val Ala Asn Arg Leu Pro Val Asp Met Thr Val His
10 15 20

```
cca gat ggt agc tat agc atc tcc ccc agc ccc ggt ggc ctt gtc acg   211
Pro Asp Gly Ser Tyr Ser Ile Ser Pro Ser Pro Gly Gly Leu Val Thr
            25                  30                  35

ggg ctt tcc ccc gtt ctg gaa caa cat cgt gga tgt tgg gtc gga tgg   259
Gly Leu Ser Pro Val Leu Glu Gln His Arg Gly Cys Trp Val Gly Trp
        40                  45                  50

cct gga act gta gat gtt gca ccc gaa cca ttt cga aca gat acg ggt   307
Pro Gly Thr Val Asp Val Ala Pro Glu Pro Phe Arg Thr Asp Thr Gly
    55                  60                  65

gtt ttg ctg cac cct gtt gtc ctc act gca agt gac tat gaa ggc ttc   355
Val Leu Leu His Pro Val Val Leu Thr Ala Ser Asp Tyr Glu Gly Phe
70                  75                  80                  85

tac gag ggc ttt tca aac gca acg ctg tgg cct ctt ttc cac gat ctg   403
Tyr Glu Gly Phe Ser Asn Ala Thr Leu Trp Pro Leu Phe His Asp Leu
                90                  95                  100

att gtt act ccg gtg tac aac acc gat tgg tgg cat gcg ttt cgg gag   451
Ile Val Thr Pro Val Tyr Asn Thr Asp Trp Trp His Ala Phe Arg Glu
            105                 110                 115

gta aac ctc aag ttc gct gaa gcc gtg agc caa gtg gcg gca cac ggt   499
Val Asn Leu Lys Phe Ala Glu Ala Val Ser Gln Val Ala Ala His Gly
        120                 125                 130

gcc act gtg tgg gtg cag gac tat cag ctg ttg ctg gtt cct ggc att   547
Ala Thr Val Trp Val Gln Asp Tyr Gln Leu Leu Leu Val Pro Gly Ile
    135                 140                 145

ttg cgc cag atg cgc cct gat ttg aag atc ggt ttc ttc ctc cac att   595
Leu Arg Gln Met Arg Pro Asp Leu Lys Ile Gly Phe Phe Leu His Ile
150                 155                 160                 165

ccc ttc cct tcc cct gat ctg ttc cgt cag ctg ccg tgg cgt gaa gag   643
Pro Phe Pro Ser Pro Asp Leu Phe Arg Gln Leu Pro Trp Arg Glu Glu
                170                 175                 180

att gtt cga ggc atg ctg ggc gca gat ttg gtg gga ttc cat ttg gtt   691
Ile Val Arg Gly Met Leu Gly Ala Asp Leu Val Gly Phe His Leu Val
            185                 190                 195

caa aac gca gaa aac ttc ctt gcg tta acc cag cag gtt gcc ggc act   739
Gln Asn Ala Glu Asn Phe Leu Ala Leu Thr Gln Gln Val Ala Gly Thr
        200                 205                 210

gcc ggg tct cat gtg ggt cag ccg gac acc ttg cag gtc agt ggt gaa   787
Ala Gly Ser His Val Gly Gln Pro Asp Thr Leu Gln Val Ser Gly Glu
    215                 220                 225
```

```
gca ttg gtg cgt gag att ggc gct cat gtt gaa acc gct gac gga agg   835
Ala Leu Val Arg Glu Ile Gly Ala His Val Glu Thr Ala Asp Gly Arg
230                 235                 240                 245

cga gtt agc gtc ggg gcg ttc ccg atc tcg att gat gtt gaa atg ttt   883
Arg Val Ser Val Gly Ala Phe Pro Ile Ser Ile Asp Val Glu Met Phe
                250                 255                 260

ggg gag gcg tcg aaa agc gcc gtt ctt gat ctt tta aaa acg ctc gac   931
Gly Glu Ala Ser Lys Ser Ala Val Leu Asp Leu Leu Lys Thr Leu Asp
            265                 270                 275

gag ccg gaa acc gta ttc ctg ggc gtt gac cga ctg gac tac acc aag   979
Glu Pro Glu Thr Val Phe Leu Gly Val Asp Arg Leu Asp Tyr Thr Lys
        280                 285                 290

ggc att ttg cag cgc ctg ctt gcg ttt gag gaa ctg ctg gaa tcc ggc   1027
Gly Ile Leu Gln Arg Leu Leu Ala Phe Glu Glu Leu Leu Glu Ser Gly
    295                 300                 305

gcg ttg gag gcc gac aaa gct gtg ttg ctg cag gtc gcg acg cct tcg   1075
Ala Leu Glu Ala Asp Lys Ala Val Leu Leu Gln Val Ala Thr Pro Ser
310                 315                 320                 325

cgt gag cgc att gat cac tat cgt gtg tcg cgt tcg cag gtc gag gaa   1123
Arg Glu Arg Ile Asp His Tyr Arg Val Ser Arg Ser Gln Val Glu Glu
                330                 335                 340

gcc gtc ggc cgt atc aat ggt cgt ttc ggt cgc atg ggg cgt ccc gtg   1171
Ala Val Gly Arg Ile Asn Gly Arg Phe Gly Arg Met Gly Arg Pro Val
            345                 350                 355

gtg cat tat cta cac agg tca ttg agc aaa aat gat ctc cag gtg ctg   1219
Val His Tyr Leu His Arg Ser Leu Ser Lys Asn Asp Leu Gln Val Leu
        360                 365                 370

tat acc gca gcc gat gtc atg ctg gtt acg cct ttt aaa gac ggt atg   1267
Tyr Thr Ala Ala Asp Val Met Leu Val Thr Pro Phe Lys Asp Gly Met
    375                 380                 385

aac ttg gtg gct aaa gaa ttc gtg gcc aac cac cgc gac ggc act ggt   1315
Asn Leu Val Ala Lys Glu Phe Val Ala Asn His Arg Asp Gly Thr Gly
390                 395                 400                 405

gct ttg gtg ctg tcc gaa ttt gcc ggc gcg gcc act gag ctg acc ggt   1363
Ala Leu Val Leu Ser Glu Phe Ala Gly Ala Ala Thr Glu Leu Thr Gly
                410                 415                 420

gcg tat tta tgc aac cca ttt gat gtg gaa tcc atc aaa cgg caa atg   1411
Ala Tyr Leu Cys Asn Pro Phe Asp Val Glu Ser Ile Lys Arg Gln Met
            425                 430                 435
```

```
gtg gca gct gtc cat gat ttg aag cac aat ccg gaa tct gcg gca acg   1459
Val Ala Ala Val His Asp Leu Lys His Asn Pro Glu Ser Ala Ala Thr
        440                 445                 450

cga atg aaa acg aac agc gag cag gtc tat acc cac gac gtc aac gtg   1507
Arg Met Lys Thr Asn Ser Glu Gln Val Tyr Thr His Asp Val Asn Val
    455                 460                 465

tgg gct aat agt ttc ctg gat tgt ttg gca cag tcg gga gaa aac tca   1555
Trp Ala Asn Ser Phe Leu Asp Cys Leu Ala Gln Ser Gly Glu Asn Ser
470                 475                 480                 485

tgaaccgcgc acgaatcgcg acc                                        1578
```

<210> 30
<211> 485
<212> PRT
<213> Corynebacterium glutamicum

<400> 30

```
Met Asp Asp Ser Asn Ser Phe Val Val Val Ala Asn Arg Leu Pro Val
  1               5                  10                  15

Asp Met Thr Val His Pro Asp Gly Ser Tyr Ser Ile Ser Pro Ser Pro
             20                  25                  30

Gly Gly Leu Val Thr Gly Leu Ser Pro Val Leu Glu Gln His Arg Gly
         35                  40                  45

Cys Trp Val Gly Trp Pro Gly Thr Val Asp Val Ala Pro Glu Pro Phe
     50                  55                  60

Arg Thr Asp Thr Gly Val Leu Leu His Pro Val Val Leu Thr Ala Ser
 65                  70                  75                  80

Asp Tyr Glu Gly Phe Tyr Glu Gly Phe Ser Asn Ala Thr Leu Trp Pro
                 85                  90                  95

Leu Phe His Asp Leu Ile Val Thr Pro Val Tyr Asn Thr Asp Trp Trp
            100                 105                 110

His Ala Phe Arg Glu Val Asn Leu Lys Phe Ala Glu Ala Val Ser Gln
        115                 120                 125

Val Ala Ala His Gly Ala Thr Val Trp Val Gln Asp Tyr Gln Leu Leu
    130                 135                 140

Leu Val Pro Gly Ile Leu Arg Gln Met Arg Pro Asp Leu Lys Ile Gly
145                 150                 155                 160

Phe Phe Leu His Ile Pro Phe Pro Ser Pro Asp Leu Phe Arg Gln Leu
                165                 170                 175

Pro Trp Arg Glu Glu Ile Val Arg Gly Met Leu Gly Ala Asp Leu Val
            180                 185                 190
```

```
Gly Phe His Leu Val Gln Asn Ala Glu Asn Phe Leu Ala Leu Thr Gln
        195                 200                 205

Gln Val Ala Gly Thr Ala Gly Ser His Val Gly Gln Pro Asp Thr Leu
    210                 215                 220

Gln Val Ser Gly Glu Ala Leu Val Arg Glu Ile Gly Ala His Val Glu
225                 230                 235                 240

Thr Ala Asp Gly Arg Arg Val Ser Val Gly Ala Phe Pro Ile Ser Ile
                245                 250                 255

Asp Val Glu Met Phe Gly Glu Ala Ser Lys Ser Ala Val Leu Asp Leu
            260                 265                 270

Leu Lys Thr Leu Asp Glu Pro Glu Thr Val Phe Leu Gly Val Asp Arg
        275                 280                 285

Leu Asp Tyr Thr Lys Gly Ile Leu Gln Arg Leu Leu Ala Phe Glu Glu
    290                 295                 300

Leu Leu Glu Ser Gly Ala Leu Glu Ala Asp Lys Ala Val Leu Leu Gln
305                 310                 315                 320

Val Ala Thr Pro Ser Arg Glu Arg Ile Asp His Tyr Arg Val Ser Arg
                325                 330                 335

Ser Gln Val Glu Glu Ala Val Gly Arg Ile Asn Gly Arg Phe Gly Arg
            340                 345                 350

Met Gly Arg Pro Val Val His Tyr Leu His Arg Ser Leu Ser Lys Asn
        355                 360                 365

Asp Leu Gln Val Leu Tyr Thr Ala Ala Asp Val Met Leu Val Thr Pro
    370                 375                 380

Phe Lys Asp Gly Met Asn Leu Val Ala Lys Glu Phe Val Ala Asn His
385                 390                 395                 400

Arg Asp Gly Thr Gly Ala Leu Val Leu Ser Glu Phe Ala Gly Ala Ala
                405                 410                 415

Thr Glu Leu Thr Gly Ala Tyr Leu Cys Asn Pro Phe Asp Val Glu Ser
            420                 425                 430

Ile Lys Arg Gln Met Val Ala Ala Val His Asp Leu Lys His Asn Pro
        435                 440                 445

Glu Ser Ala Ala Thr Arg Met Lys Thr Asn Ser Glu Gln Val Tyr Thr
    450                 455                 460

His Asp Val Asn Val Trp Ala Asn Ser Phe Leu Asp Cys Leu Ala Gln
465                 470                 475                 480

Ser Gly Glu Asn Ser
                485
```

<210> 31
<211> 891
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(868)
<223> RXA00347

<400> 31
tcggccagca atccgcttgg tgtcctggat cgcgccgaca tcttaaggtg ccagggcttt 60

aaagtgccag gggttctgtg ggatccgtac actggttccc atg act ttg act att 115
Met Thr Leu Thr Ile
1 5

gag gaa atc gcc aag acc aaa aag ctt ttg gtt gtg tcc gat ttt gat 163
Glu Glu Ile Ala Lys Thr Lys Lys Leu Leu Val Val Ser Asp Phe Asp
10 15 20

gga acc atc gca gga ttt agc aag gac gct tac aac gtt cct atc aac 211
Gly Thr Ile Ala Gly Phe Ser Lys Asp Ala Tyr Asn Val Pro Ile Asn
25 30 35

cag aaa tcc ctc aag gcg gta aaa gac ctc tcc caa caa gca gac act 259
Gln Lys Ser Leu Lys Ala Val Lys Asp Leu Ser Gln Gln Ala Asp Thr
40 45 50

gat gtt gtc att ttg tcg gga cgt cac ctg gag gga ttg aag acg gtt 307
Asp Val Val Ile Leu Ser Gly Arg His Leu Glu Gly Leu Lys Thr Val
55 60 65

ctt gat ctt ggt cag tac gac atc acc atg gtg ggt tca cac ggt tct 355
Leu Asp Leu Gly Gln Tyr Asp Ile Thr Met Val Gly Ser His Gly Ser
70 75 80 85

gag gat tcc tcc cgc ccg cgt acc ctc act cct gaa gag gta gct cgc 403
Glu Asp Ser Ser Arg Pro Arg Thr Leu Thr Pro Glu Glu Val Ala Arg
90 95 100

ctc gcc aag att gaa gca gat ctg gaa aag atc gtc gac ggc atc gaa 451
Leu Ala Lys Ile Glu Ala Asp Leu Glu Lys Ile Val Asp Gly Ile Glu
105 110 115

ggc gca ttc gtg gag atc aag cct ttc cac cgc gtg ctg cac ttc atc 499
Gly Ala Phe Val Glu Ile Lys Pro Phe His Arg Val Leu His Phe Ile
120 125 130

cgt gtt tcc gac aag gac aaa gtc caa gga atc ctc gcc caa gca gca 547
Arg Val Ser Asp Lys Asp Lys Val Gln Gly Ile Leu Ala Gln Ala Ala
135 140 145

```
cac gta gac tct tcc ggc ctg aag gtt act aac ggc aag agc atc atc   595
His Val Asp Ser Ser Gly Leu Lys Val Thr Asn Gly Lys Ser Ile Ile
150                 155                 160                 165

gaa tac tcc atc agc tcc acc acc aag ggc acc tgg ctg aag gaa tac   643
Glu Tyr Ser Ile Ser Ser Thr Thr Lys Gly Thr Trp Leu Lys Glu Tyr
                170                 175                 180

gtt gac cgc acc gag ccc act ggt gtg att ttc ctc ggc gat gac acc   691
Val Asp Arg Thr Glu Pro Thr Gly Val Ile Phe Leu Gly Asp Asp Thr
            185                 190                 195

acc gat gag cac ggt ttc aaa gct tta gaa aac gat gat cgt gcc cta   739
Thr Asp Glu His Gly Phe Lys Ala Leu Glu Asn Asp Asp Arg Ala Leu
        200                 205                 210

acc gtc aag gtt ggc gaa gga gac act gca gcc aaa acc cgc gtc gac   787
Thr Val Lys Val Gly Glu Gly Asp Thr Ala Ala Lys Thr Arg Val Asp
    215                 220                 225

gat gtt gat aat gtg gga att ttc cta gag aaa ctc gcc tac cac cgc   835
Asp Val Asp Asn Val Gly Ile Phe Leu Glu Lys Leu Ala Tyr His Arg
230                 235                 240                 245

atg cag tat gcg gaa agc gtg cga ttg ggg att taagagagcc taaacgcacg 888
Met Gln Tyr Ala Glu Ser Val Arg Leu Gly Ile
                250                 255

aaa                                                               891
```

<210> 32
<211> 256
<212> PRT
<213> Corynebacterium glutamicum

<400> 32

```
Met Thr Leu Thr Ile Glu Glu Ile Ala Lys Thr Lys Lys Leu Leu Val
  1               5                  10                  15

Val Ser Asp Phe Asp Gly Thr Ile Ala Gly Phe Ser Lys Asp Ala Tyr
             20                  25                  30

Asn Val Pro Ile Asn Gln Lys Ser Leu Lys Ala Val Lys Asp Leu Ser
         35                  40                  45

Gln Gln Ala Asp Thr Asp Val Val Ile Leu Ser Gly Arg His Leu Glu
     50                  55                  60

Gly Leu Lys Thr Val Leu Asp Leu Gly Gln Tyr Asp Ile Thr Met Val
 65                  70                  75                  80

Gly Ser His Gly Ser Glu Asp Ser Ser Arg Pro Arg Thr Leu Thr Pro
                 85                  90                  95
```

```
Glu Glu Val Ala Arg Leu Ala Lys Ile Glu Ala Asp Leu Glu Lys Ile
            100                 105                 110

Val Asp Gly Ile Glu Gly Ala Phe Val Glu Ile Lys Pro Phe His Arg
        115                 120                 125

Val Leu His Phe Ile Arg Val Ser Asp Lys Asp Lys Val Gln Gly Ile
    130                 135                 140

Leu Ala Gln Ala Ala His Val Asp Ser Ser Gly Leu Lys Val Thr Asn
145                 150                 155                 160

Gly Lys Ser Ile Ile Glu Tyr Ser Ile Ser Ser Thr Thr Lys Gly Thr
                165                 170                 175

Trp Leu Lys Glu Tyr Val Asp Arg Thr Glu Pro Thr Gly Val Ile Phe
            180                 185                 190

Leu Gly Asp Asp Thr Thr Asp Glu His Gly Phe Lys Ala Leu Glu Asn
        195                 200                 205

Asp Asp Arg Ala Leu Thr Val Lys Val Gly Glu Gly Asp Thr Ala Ala
    210                 215                 220

Lys Thr Arg Val Asp Asp Val Asp Asn Val Gly Ile Phe Leu Glu Lys
225                 230                 235                 240

Leu Ala Tyr His Arg Met Gln Tyr Ala Glu Ser Val Arg Leu Gly Ile
                245                 250                 255
```

<210> 33
<211> 2556
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(2533)
<223> RXN01239

<400> 33
gcacttgctg cgtaaatctt tttcccacgc cgggaatgcg tgaacactaa gatcgaggac 60

```
gtaccgcacg attttgccta acttttaagg gtgtttcatc atg gca cgt cca att   115
                                            Met Ala Arg Pro Ile
                                              1               5

tcc gca acg tac agg ctt caa atg cga gga cct caa gca gat agc gcc   163
Ser Ala Thr Tyr Arg Leu Gln Met Arg Gly Pro Gln Ala Asp Ser Ala
                10                  15                  20

ggg cgt tca ttt ggt ttt gcg cag gcc aaa gcc cag ctt ccc tat ctg   211
Gly Arg Ser Phe Gly Phe Ala Gln Ala Lys Ala Gln Leu Pro Tyr Leu
            25                  30                  35
```

```
aag aag cta ggc atc agc cac ctg tac ctc tcc cct att ttt acg gcc   259
Lys Lys Leu Gly Ile Ser His Leu Tyr Leu Ser Pro Ile Phe Thr Ala
         40                  45                  50

atg cca gat tcc aat cat ggc tac gat gtc att gat ccc acc acc atc   307
Met Pro Asp Ser Asn His Gly Tyr Asp Val Ile Asp Pro Thr Thr Ile
     55                  60                  65

aat gaa gag ctc ggt ggc atg gag ggt ctt cga gat ctt gcc gca gct   355
Asn Glu Glu Leu Gly Gly Met Glu Gly Leu Arg Asp Leu Ala Ala Ala
 70                  75                  80                  85

aca cac gag ttg ggc atg ggc atc atc att gat att gtt ccc aac cat   403
Thr His Glu Leu Gly Met Gly Ile Ile Ile Asp Ile Val Pro Asn His
                 90                  95                 100

tta ggt gtt gcc gtt cca cat ttg aat cct tgg tgg tgg gat gtt cta   451
Leu Gly Val Ala Val Pro His Leu Asn Pro Trp Trp Trp Asp Val Leu
            105                 110                 115

aaa aac ggc aaa gat tcc gct ttt gag ttc tat ttc gat att gac tgg   499
Lys Asn Gly Lys Asp Ser Ala Phe Glu Phe Tyr Phe Asp Ile Asp Trp
        120                 125                 130

cac gaa gac aac ggt tct ggt ggc aag ctg ggc atg ccg att ctg ggt   547
His Glu Asp Asn Gly Ser Gly Gly Lys Leu Gly Met Pro Ile Leu Gly
    135                 140                 145

gct gaa ggc gat gaa gac aag ctg gaa ttc gcg gag ctt gat gga gag   595
Ala Glu Gly Asp Glu Asp Lys Leu Glu Phe Ala Glu Leu Asp Gly Glu
150                 155                 160                 165

aaa gtg ctc aaa tat ttt gac cac ctc ttc cca atc gcg cct ggt acc   643
Lys Val Leu Lys Tyr Phe Asp His Leu Phe Pro Ile Ala Pro Gly Thr
                170                 175                 180

gaa gaa ggg aca ccg caa gaa gtc tac aag cgc cag cat tac cgc ctg   691
Glu Glu Gly Thr Pro Gln Glu Val Tyr Lys Arg Gln His Tyr Arg Leu
            185                 190                 195

cag ttc tgg cgc gat ggc gtg atc aac ttc cgt cgc ttc ttt tcc gtg   739
Gln Phe Trp Arg Asp Gly Val Ile Asn Phe Arg Arg Phe Phe Ser Val
        200                 205                 210

aat acg ttg gct ggc atc agg caa gaa gat ccc tta gtg ttt gaa cat   787
Asn Thr Leu Ala Gly Ile Arg Gln Glu Asp Pro Leu Val Phe Glu His
    215                 220                 225

act cat cgt ctg ctg cgc gaa ttg gtg gcg gaa gac ctc att gac ggc   835
Thr His Arg Leu Leu Arg Glu Leu Val Ala Glu Asp Leu Ile Asp Gly
230                 235                 240                 245
```

```
gtg cgc gtc gat cac ccc gac ggg ctt tcc gat cct ttt gga tat ctg   883
Val Arg Val Asp His Pro Asp Gly Leu Ser Asp Pro Phe Gly Tyr Leu
                250                 255                 260

cac aga ctc cgc gac ctc att gga cct gac cgc tgg ctg atc atc gaa   931
His Arg Leu Arg Asp Leu Ile Gly Pro Asp Arg Trp Leu Ile Ile Glu
            265                 270                 275

aag atc ttg agc gtt gat gaa cca ctc gat ccc cgc ctg gcc gtt gat   979
Lys Ile Leu Ser Val Asp Glu Pro Leu Asp Pro Arg Leu Ala Val Asp
        280                 285                 290

ggc acc act ggc tac gac gcc ctc cgt gaa ctc gac ggc gtg ttt atc   1027
Gly Thr Thr Gly Tyr Asp Ala Leu Arg Glu Leu Asp Gly Val Phe Ile
    295                 300                 305

tcc cga gaa tct gag gac aaa ttc tcc atg ctg gcg ctg acc cac agt   1075
Ser Arg Glu Ser Glu Asp Lys Phe Ser Met Leu Ala Leu Thr His Ser
310                 315                 320                 325

gga tcc acc tgg gat gaa cgc gcc ctc aaa tcc acg gag gaa agc ctc   1123
Gly Ser Thr Trp Asp Glu Arg Ala Leu Lys Ser Thr Glu Glu Ser Leu
                330                 335                 340

aaa cga gtc gtc gcc caa caa gaa ctc gca gcc gaa atc tta agg ctc   1171
Lys Arg Val Val Ala Gln Gln Glu Leu Ala Ala Glu Ile Leu Arg Leu
            345                 350                 355

gcc cgc gcc atg cgc cgc gat aac ttc tcc acc gca ggc acc aac gtc   1219
Ala Arg Ala Met Arg Arg Asp Asn Phe Ser Thr Ala Gly Thr Asn Val
        360                 365                 370

acc gaa gac aaa ctt agc gaa acc atc atc gaa tta gtc gcc gcc atg   1267
Thr Glu Asp Lys Leu Ser Glu Thr Ile Ile Glu Leu Val Ala Ala Met
    375                 380                 385

ccc gtc tac cgc gcc gac tac atc tcc ctc tca cgc acc acc gcc acc   1315
Pro Val Tyr Arg Ala Asp Tyr Ile Ser Leu Ser Arg Thr Thr Ala Thr
390                 395                 400                 405

gtc atc gcg gag atg tcc aaa cgc ttc ccc tcc cgg cgt gac gca ctc   1363
Val Ile Ala Glu Met Ser Lys Arg Phe Pro Ser Arg Arg Asp Ala Leu
                410                 415                 420

gac ctc atc gcg gcc gcc cta ctt ggc aat ggc gag gcc aaa atc cgc   1411
Asp Leu Ile Ala Ala Ala Leu Leu Gly Asn Gly Glu Ala Lys Ile Arg
            425                 430                 435

ttc gct caa gtc tgc ggc gcc gtc atg gct aaa ggt gtg gaa gac acc   1459
Phe Ala Gln Val Cys Gly Ala Val Met Ala Lys Gly Val Glu Asp Thr
        440                 445                 450
```

```
acc ttc tac cgc gca tct agg ctc gtt gca ttg caa gaa gtc ggt ggc   1507
Thr Phe Tyr Arg Ala Ser Arg Leu Val Ala Leu Gln Glu Val Gly Gly
    455                 460                 465

gcg ccg ggg aga ttc ggc gtc tcc gct gca gaa ttc cac ttg ctg cag   1555
Ala Pro Gly Arg Phe Gly Val Ser Ala Ala Glu Phe His Leu Leu Gln
470                 475                 480                 485

gaa gaa cgc agc ctg ctg tgg cca cgc acc atg acc acc ttg tcc acg   1603
Glu Glu Arg Ser Leu Leu Trp Pro Arg Thr Met Thr Thr Leu Ser Thr
                490                 495                 500

cat gac acc aaa cgt ggc gaa gat acc cgc gcc cgc atc atc tcc ctg   1651
His Asp Thr Lys Arg Gly Glu Asp Thr Arg Ala Arg Ile Ile Ser Leu
            505                 510                 515

tct gaa gtc ccc gat atg tac tcc gag ctg gtc aat cgt gtt ttc gcg   1699
Ser Glu Val Pro Asp Met Tyr Ser Glu Leu Val Asn Arg Val Phe Ala
        520                 525                 530

gtg ctc ccc gcg cca gac ggc gca acg ggc agt ttc ctc cta caa aac   1747
Val Leu Pro Ala Pro Asp Gly Ala Thr Gly Ser Phe Leu Leu Gln Asn
    535                 540                 545

ctg ctg ggc gta tgg ccc gcc gac ggc gtg atc acc gat gcg ctg cgc   1795
Leu Leu Gly Val Trp Pro Ala Asp Gly Val Ile Thr Asp Ala Leu Arg
550                 555                 560                 565

gat cga ttc agg gaa tac gcc cta aaa gct atc cgc gaa gca tcc aca   1843
Asp Arg Phe Arg Glu Tyr Ala Leu Lys Ala Ile Arg Glu Ala Ser Thr
                570                 575                 580

aaa acc acg tgg gtg gac ccc aac gag tcc ttc gag gct gcg gtc tgc   1891
Lys Thr Thr Trp Val Asp Pro Asn Glu Ser Phe Glu Ala Ala Val Cys
            585                 590                 595

gat tgg gtg gaa gcg ctt ttc gac gga ccc tcc acc tca cta atc acc   1939
Asp Trp Val Glu Ala Leu Phe Asp Gly Pro Ser Thr Ser Leu Ile Thr
        600                 605                 610

gaa ttt gtc tcc cac atc aac cgt ggc tct gtg caa atc tcc tta ggc   1987
Glu Phe Val Ser His Ile Asn Arg Gly Ser Val Gln Ile Ser Leu Gly
    615                 620                 625

agg aaa ctg ctg caa atg gtg ggc gct gga atc ccc gac act tac caa   2035
Arg Lys Leu Leu Gln Met Val Gly Ala Gly Ile Pro Asp Thr Tyr Gln
630                 635                 640                 645

gga act gag ttt tta gaa gac tcc ctg gta gat ccc gat aac cga cgc   2083
Gly Thr Glu Phe Leu Glu Asp Ser Leu Val Asp Pro Asp Asn Arg Arg
                650                 655                 660
```

```
ttt gtt gat tac acc gcc aga gaa caa gtc ctg gag cgc ctg caa acc   2131
Phe Val Asp Tyr Thr Ala Arg Glu Gln Val Leu Glu Arg Leu Gln Thr
            665                 670                 675

tgg gct tgg acg cag gtt aat tcg gta gaa gac ttg gtg gat aac gcc   2179
Trp Ala Trp Thr Gln Val Asn Ser Val Glu Asp Leu Val Asp Asn Ala
        680                 685                 690

gac atc gcc aaa atg gcc gtg gtc cat aaa tcc ctc gag ttg cgt gct   2227
Asp Ile Ala Lys Met Ala Val Val His Lys Ser Leu Glu Leu Arg Ala
    695                 700                 705

gaa ttt cgt gca agc ttt gtt ggt gga gat cat cag gca gta ttt ggc   2275
Glu Phe Arg Ala Ser Phe Val Gly Gly Asp His Gln Ala Val Phe Gly
710                 715                 720                 725

gaa ggt cgc gca gaa tcc cac atc atg ggc atc gcc cgc ggt aca gac   2323
Glu Gly Arg Ala Glu Ser His Ile Met Gly Ile Ala Arg Gly Thr Asp
                730                 735                 740

cga aac cac ctc aac atc att gct ctt gct acc cgt cga cca ctg atc   2371
Arg Asn His Leu Asn Ile Ile Ala Leu Ala Thr Arg Arg Pro Leu Ile
            745                 750                 755

ttg gaa gac cgt ggc gga tgg tat gac acc acc gtc acg ctt cct ggt   2419
Leu Glu Asp Arg Gly Gly Trp Tyr Asp Thr Thr Val Thr Leu Pro Gly
        760                 765                 770

gga caa tgg gaa gac agg ctc acc ggg caa cgc ttc agt ggt gtt gtc   2467
Gly Gln Trp Glu Asp Arg Leu Thr Gly Gln Arg Phe Ser Gly Val Val
    775                 780                 785

cca gcc acc gat ttg ttc tca cat cta ccc gta tct ttg ttg gtt tta   2515
Pro Ala Thr Asp Leu Phe Ser His Leu Pro Val Ser Leu Leu Val Leu
790                 795                 800                 805

gta ccc gat agt gag ttt tgatccctgc acaggaaagt tag                2556
Val Pro Asp Ser Glu Phe
                810
```

<210> 34
<211> 811
<212> PRT
<213> Corynebacterium glutamicum

<400> 34

```
Met Ala Arg Pro Ile Ser Ala Thr Tyr Arg Leu Gln Met Arg Gly Pro
  1               5                  10                  15

Gln Ala Asp Ser Ala Gly Arg Ser Phe Gly Phe Ala Gln Ala Lys Ala
             20                  25                  30

Gln Leu Pro Tyr Leu Lys Lys Leu Gly Ile Ser His Leu Tyr Leu Ser
         35                  40                  45
```

```
Pro Ile Phe Thr Ala Met Pro Asp Ser Asn His Gly Tyr Asp Val Ile
     50                  55                  60

Asp Pro Thr Thr Ile Asn Glu Glu Leu Gly Gly Met Glu Gly Leu Arg
 65                  70                  75                  80

Asp Leu Ala Ala Ala Thr His Glu Leu Gly Met Gly Ile Ile Ile Asp
                 85                  90                  95

Ile Val Pro Asn His Leu Gly Val Ala Val Pro His Leu Asn Pro Trp
            100                 105                 110

Trp Trp Asp Val Leu Lys Asn Gly Lys Asp Ser Ala Phe Glu Phe Tyr
        115                 120                 125

Phe Asp Ile Asp Trp His Glu Asp Asn Gly Ser Gly Gly Lys Leu Gly
    130                 135                 140

Met Pro Ile Leu Gly Ala Glu Gly Asp Glu Asp Lys Leu Glu Phe Ala
145                 150                 155                 160

Glu Leu Asp Gly Glu Lys Val Leu Lys Tyr Phe Asp His Leu Phe Pro
                165                 170                 175

Ile Ala Pro Gly Thr Glu Glu Gly Thr Pro Gln Glu Val Tyr Lys Arg
            180                 185                 190

Gln His Tyr Arg Leu Gln Phe Trp Arg Asp Gly Val Ile Asn Phe Arg
        195                 200                 205

Arg Phe Phe Ser Val Asn Thr Leu Ala Gly Ile Arg Gln Glu Asp Pro
    210                 215                 220

Leu Val Phe Glu His Thr His Arg Leu Leu Arg Glu Leu Val Ala Glu
225                 230                 235                 240

Asp Leu Ile Asp Gly Val Arg Val Asp His Pro Asp Gly Leu Ser Asp
                245                 250                 255

Pro Phe Gly Tyr Leu His Arg Leu Arg Asp Leu Ile Gly Pro Asp Arg
            260                 265                 270

Trp Leu Ile Ile Glu Lys Ile Leu Ser Val Asp Glu Pro Leu Asp Pro
        275                 280                 285

Arg Leu Ala Val Asp Gly Thr Thr Gly Tyr Asp Ala Leu Arg Glu Leu
    290                 295                 300

Asp Gly Val Phe Ile Ser Arg Glu Ser Glu Asp Lys Phe Ser Met Leu
305                 310                 315                 320

Ala Leu Thr His Ser Gly Ser Thr Trp Asp Glu Arg Ala Leu Lys Ser
                325                 330                 335

Thr Glu Glu Ser Leu Lys Arg Val Val Ala Gln Gln Glu Leu Ala Ala
            340                 345                 350
```

Glu Ile Leu Arg Leu Ala Arg Ala Met Arg Arg Asp Asn Phe Ser Thr
355 360 365

Ala Gly Thr Asn Val Thr Glu Asp Lys Leu Ser Glu Thr Ile Ile Glu
370 375 380

Leu Val Ala Ala Met Pro Val Tyr Arg Ala Asp Tyr Ile Ser Leu Ser
385 390 395 400

Arg Thr Thr Ala Thr Val Ile Ala Glu Met Ser Lys Arg Phe Pro Ser
405 410 415

Arg Arg Asp Ala Leu Asp Leu Ile Ala Ala Ala Leu Leu Gly Asn Gly
420 425 430

Glu Ala Lys Ile Arg Phe Ala Gln Val Cys Gly Ala Val Met Ala Lys
435 440 445

Gly Val Glu Asp Thr Thr Phe Tyr Arg Ala Ser Arg Leu Val Ala Leu
450 455 460

Gln Glu Val Gly Gly Ala Pro Gly Arg Phe Gly Val Ser Ala Ala Glu
465 470 475 480

Phe His Leu Leu Gln Glu Glu Arg Ser Leu Leu Trp Pro Arg Thr Met
485 490 495

Thr Thr Leu Ser Thr His Asp Thr Lys Arg Gly Glu Asp Thr Arg Ala
500 505 510

Arg Ile Ile Ser Leu Ser Glu Val Pro Asp Met Tyr Ser Glu Leu Val
515 520 525

Asn Arg Val Phe Ala Val Leu Pro Ala Pro Asp Gly Ala Thr Gly Ser
530 535 540

Phe Leu Leu Gln Asn Leu Leu Gly Val Trp Pro Ala Asp Gly Val Ile
545 550 555 560

Thr Asp Ala Leu Arg Asp Arg Phe Arg Glu Tyr Ala Leu Lys Ala Ile
565 570 575

Arg Glu Ala Ser Thr Lys Thr Thr Trp Val Asp Pro Asn Glu Ser Phe
580 585 590

Glu Ala Ala Val Cys Asp Trp Val Glu Ala Leu Phe Asp Gly Pro Ser
595 600 605

Thr Ser Leu Ile Thr Glu Phe Val Ser His Ile Asn Arg Gly Ser Val
610 615 620

Gln Ile Ser Leu Gly Arg Lys Leu Leu Gln Met Val Gly Ala Gly Ile
625 630 635 640

Pro Asp Thr Tyr Gln Gly Thr Glu Phe Leu Glu Asp Ser Leu Val Asp
645 650 655

```
Pro Asp Asn Arg Arg Phe Val Asp Tyr Thr Ala Arg Glu Gln Val Leu
            660                 665                 670

Glu Arg Leu Gln Thr Trp Ala Trp Thr Gln Val Asn Ser Val Glu Asp
        675                 680                 685

Leu Val Asp Asn Ala Asp Ile Ala Lys Met Ala Val Val His Lys Ser
    690                 695                 700

Leu Glu Leu Arg Ala Glu Phe Arg Ala Ser Phe Val Gly Gly Asp His
705                 710                 715                 720

Gln Ala Val Phe Gly Glu Gly Arg Ala Glu Ser His Ile Met Gly Ile
                725                 730                 735

Ala Arg Gly Thr Asp Arg Asn His Leu Asn Ile Ile Ala Leu Ala Thr
            740                 745                 750

Arg Arg Pro Leu Ile Leu Glu Asp Arg Gly Gly Trp Tyr Asp Thr Thr
        755                 760                 765

Val Thr Leu Pro Gly Gly Gln Trp Glu Asp Arg Leu Thr Gly Gln Arg
    770                 775                 780

Phe Ser Gly Val Val Pro Ala Thr Asp Leu Phe Ser His Leu Pro Val
785                 790                 795                 800

Ser Leu Leu Val Leu Val Pro Asp Ser Glu Phe
                805                 810
```

<210> 35
<211> 1953
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(1930)
<223> RXA02645

<400> 35

```
gatacagctc cttgatggag tgaataaatt cgcgagcctg ctcctgatct tgcacacgcg 60

tgatataggt cagaaatcgc gagcgcttga tctctagttc atg ctc aaa gac ttg   115
                                            Met Leu Lys Asp Leu
                                              1               5

acc ggc ctg agg gag ttg gta ttg cgt gag atg tgc cat agc atc tca   163
Thr Gly Leu Arg Glu Leu Val Leu Arg Glu Met Cys His Ser Ile Ser
                  10                  15                  20

cat ctt agc tcg cca acc ggc agc att ttc act agc ctg gtg gcc atg   211
His Leu Ser Ser Pro Thr Gly Ser Ile Phe Thr Ser Leu Val Ala Met
            25                  30                  35
```

```
ttg acc tcg caa agc ttt tca gtg tgg gct cca ctt ccc cac gat gta   259
Leu Thr Ser Gln Ser Phe Ser Val Trp Ala Pro Leu Pro His Asp Val
        40                  45                  50

cat ctg atc ctc aac ggc gaa acc ctc ccc atg cac aaa acg gag ggc   307
His Leu Ile Leu Asn Gly Glu Thr Leu Pro Met His Lys Thr Glu Gly
    55                  60                  65

agc tgg tgg cgc gcc gag atc gcg ccc aag gcc ggc gat cgt tac ggt   355
Ser Trp Trp Arg Ala Glu Ile Ala Pro Lys Ala Gly Asp Arg Tyr Gly
70                  75                  80                  85

ttt tcg ctt ttc gac ggc tcc tcc tgg tca aaa acc ctc ccc gat ccc   403
Phe Ser Leu Phe Asp Gly Ser Ser Trp Ser Lys Thr Leu Pro Asp Pro
                90                  95                  100

cgc tcc aca tct caa cca gac ggg gtt cat ggt tta agt gaa gtc tcc   451
Arg Ser Thr Ser Gln Pro Asp Gly Val His Gly Leu Ser Glu Val Ser
            105                 110                 115

gat gat tcc tat ctg tgg ggt gac cag cag tgg act ggc cga att ctc   499
Asp Asp Ser Tyr Leu Trp Gly Asp Gln Gln Trp Thr Gly Arg Ile Leu
        120                 125                 130

cct ggc tcg gtg tta tat gag ctg cat gtg ggc acc ttt agt gaa gat   547
Pro Gly Ser Val Leu Tyr Glu Leu His Val Gly Thr Phe Ser Glu Asp
    135                 140                 145

gga acg ttt gag gga gtc gtc gac aag ctt cct tat ctg cgc gac ctc   595
Gly Thr Phe Glu Gly Val Val Asp Lys Leu Pro Tyr Leu Arg Asp Leu
150                 155                 160                 165

ggc gtg acc gcc atc gaa ctt tta ccc gtg cag ccc ttt ggc ggc aac   643
Gly Val Thr Ala Ile Glu Leu Leu Pro Val Gln Pro Phe Gly Gly Asn
                170                 175                 180

cgc aat tgg ggc tac gac ggg gtg ctg tgg cac gcc gtc cat gca ggc   691
Arg Asn Trp Gly Tyr Asp Gly Val Leu Trp His Ala Val His Ala Gly
            185                 190                 195

tac ggc ggt ccg gcg ggc ttg aaa aag ctt atc gac gcc tcc cac cag   739
Tyr Gly Gly Pro Ala Gly Leu Lys Lys Leu Ile Asp Ala Ser His Gln
        200                 205                 210

gcc ggc atc gcc gtc tac tta gac gtc gtg tac aac cac ttc ggc ccc   787
Ala Gly Ile Ala Val Tyr Leu Asp Val Val Tyr Asn His Phe Gly Pro
    215                 220                 225

gac ggc aac tac aac ggg caa ttt ggc ccc tac acc tct ggc ggc agc   835
Asp Gly Asn Tyr Asn Gly Gln Phe Gly Pro Tyr Thr Ser Gly Gly Ser
230                 235                 240                 245
```

```
acc ggc tgg ggc gac gtg gtc aac atc aac ggc cat gat tca gat gaa   883
Thr Gly Trp Gly Asp Val Val Asn Ile Asn Gly His Asp Ser Asp Glu
                250                 255                 260

gtc cgc aat tat att ctc gac gcc gca cgc cag tgg ttc gaa gat ttt   931
Val Arg Asn Tyr Ile Leu Asp Ala Ala Arg Gln Trp Phe Glu Asp Phe
            265                 270                 275

cac gtt gat ggg ctc cgc ctc gat gcg gtg cat tct ctc gat gat cgc   979
His Val Asp Gly Leu Arg Leu Asp Ala Val His Ser Leu Asp Asp Arg
        280                 285                 290

ggc gcc tat tcc cta ctt gcg cag ctg acc atg gtg gcc gag gat gtc   1027
Gly Ala Tyr Ser Leu Leu Ala Gln Leu Thr Met Val Ala Glu Asp Val
    295                 300                 305

tcc gca caa aca ggc atc cca cgc tca ttg att gca gaa tct gaa ctc   1075
Ser Ala Gln Thr Gly Ile Pro Arg Ser Leu Ile Ala Glu Ser Glu Leu
310                 315                 320                 325

aat gac ccc aag ttc gtt acc tcc cgc gag gcc ggc ggt ttt ggc ctg   1123
Asn Asp Pro Lys Phe Val Thr Ser Arg Glu Ala Gly Gly Phe Gly Leu
                330                 335                 340

gat gca cag tgg gtt gac gat atc cac cac gcc ctc cat gcc ctc gtt   1171
Asp Ala Gln Trp Val Asp Asp Ile His His Ala Leu His Ala Leu Val
            345                 350                 355

tct ggc gaa cgc aat ggt tat tac agc gat ttc gga tct gtc gac aca   1219
Ser Gly Glu Arg Asn Gly Tyr Tyr Ser Asp Phe Gly Ser Val Asp Thr
        360                 365                 370

tta gcc aaa acc ctg cgt gaa gta ttt gaa cac acc gga aac tac tcc   1267
Leu Ala Lys Thr Leu Arg Glu Val Phe Glu His Thr Gly Asn Tyr Ser
    375                 380                 385

acg tac cgc gga cgc aac cac ggc cgc cct gtg cac ccc gat atc acc   1315
Thr Tyr Arg Gly Arg Asn His Gly Arg Pro Val His Pro Asp Ile Thr
390                 395                 400                 405

cct gcc tcg cgc ttt gtc acc tac acc acc acc cat gat cag acc ggc   1363
Pro Ala Ser Arg Phe Val Thr Tyr Thr Thr Thr His Asp Gln Thr Gly
                410                 415                 420

aac cgc gca atc ggc gac cgt cct tcc acg act ctc acc ccg gaa cag   1411
Asn Arg Ala Ile Gly Asp Arg Pro Ser Thr Thr Leu Thr Pro Glu Gln
            425                 430                 435

cag gtg ttg aag gca gcc att atc tac agc tcg ccg tat acc ccg atg   1459
Gln Val Leu Lys Ala Ala Ile Ile Tyr Ser Ser Pro Tyr Thr Pro Met
        440                 445                 450
```

```
ttg ttt atg ggt gaa gaa ttc gga gcc acc acc cca ttc gcc ttc ttt   1507
Leu Phe Met Gly Glu Glu Phe Gly Ala Thr Thr Pro Phe Ala Phe Phe
    455                 460                 465

tgc tcc cac acc gac ccc gag ctc aac cgg cta acc tcc gag ggc cgc   1555
Cys Ser His Thr Asp Pro Glu Leu Asn Arg Leu Thr Ser Glu Gly Arg
470                 475                 480                 485

aaa cgg gaa ttc gca cgc ctt ggc tgg aac gcc gac gac atc ccc tcc   1603
Lys Arg Glu Phe Ala Arg Leu Gly Trp Asn Ala Asp Asp Ile Pro Ser
                490                 495                 500

ccc gag ctg gaa tcc acc ttc acc tcc tcc aaa ctc gat tgg gag ttc   1651
Pro Glu Leu Glu Ser Thr Phe Thr Ser Ser Lys Leu Asp Trp Glu Phe
            505                 510                 515

act gcg gag cag cgc cgc atc aac gac gct tac aag cag ctg ttg cac   1699
Thr Ala Glu Gln Arg Arg Ile Asn Asp Ala Tyr Lys Gln Leu Leu His
        520                 525                 530

ctg cgg cac acc ttg ggc ttc tcc caa cca aac ttg ctc aca ctc gag   1747
Leu Arg His Thr Leu Gly Phe Ser Gln Pro Asn Leu Leu Thr Leu Glu
    535                 540                 545

gtt gag cac ggc gag aac tgg cta tcg atg gcc aat ggt cgc ggc cga   1795
Val Glu His Gly Glu Asn Trp Leu Ser Met Ala Asn Gly Arg Gly Arg
550                 555                 560                 565

att ctg gcg aat ttc tcc gac gac acc atc acc gtc ccg ctt ggc ggc   1843
Ile Leu Ala Asn Phe Ser Asp Asp Thr Ile Thr Val Pro Leu Gly Gly
                570                 575                 580

gag ctg att tac agc ttc act tcc ccc acc gtc acc gac acc tcc aca   1891
Glu Leu Ile Tyr Ser Phe Thr Ser Pro Thr Val Thr Asp Thr Ser Thr
            585                 590                 595

acc ctt cag ccg tgg ggc ttt gcg atc ctg acc cga aac tagaaaaagg    1940
Thr Leu Gln Pro Trp Gly Phe Ala Ile Leu Thr Arg Asn
        600                 605                 610

ccacctcgat tga                                                   1953
```

<210> 36
<211> 610
<212> PRT
<213> Corynebacterium glutamicum

<400> 36

```
Met Leu Lys Asp Leu Thr Gly Leu Arg Glu Leu Val Leu Arg Glu Met
  1               5                  10                  15

Cys His Ser Ile Ser His Leu Ser Ser Pro Thr Gly Ser Ile Phe Thr
             20                  25                  30
```

Ser Leu Val Ala Met Leu Thr Ser Gln Ser Phe Ser Val Trp Ala Pro
35 40 45

Leu Pro His Asp Val His Leu Ile Leu Asn Gly Glu Thr Leu Pro Met
50 55 60

His Lys Thr Glu Gly Ser Trp Trp Arg Ala Glu Ile Ala Pro Lys Ala
65 70 75 80

Gly Asp Arg Tyr Gly Phe Ser Leu Phe Asp Gly Ser Ser Trp Ser Lys
85 90 95

Thr Leu Pro Asp Pro Arg Ser Thr Ser Gln Pro Asp Gly Val His Gly
100 105 110

Leu Ser Glu Val Ser Asp Asp Ser Tyr Leu Trp Gly Asp Gln Gln Trp
115 120 125

Thr Gly Arg Ile Leu Pro Gly Ser Val Leu Tyr Glu Leu His Val Gly
130 135 140

Thr Phe Ser Glu Asp Gly Thr Phe Glu Gly Val Val Asp Lys Leu Pro
145 150 155 160

Tyr Leu Arg Asp Leu Gly Val Thr Ala Ile Glu Leu Leu Pro Val Gln
165 170 175

Pro Phe Gly Gly Asn Arg Asn Trp Gly Tyr Asp Gly Val Leu Trp His
180 185 190

Ala Val His Ala Gly Tyr Gly Gly Pro Ala Gly Leu Lys Lys Leu Ile
195 200 205

Asp Ala Ser His Gln Ala Gly Ile Ala Val Tyr Leu Asp Val Val Tyr
210 215 220

Asn His Phe Gly Pro Asp Gly Asn Tyr Asn Gly Gln Phe Gly Pro Tyr
225 230 235 240

Thr Ser Gly Gly Ser Thr Gly Trp Gly Asp Val Val Asn Ile Asn Gly
245 250 255

His Asp Ser Asp Glu Val Arg Asn Tyr Ile Leu Asp Ala Ala Arg Gln
260 265 270

Trp Phe Glu Asp Phe His Val Asp Gly Leu Arg Leu Asp Ala Val His
275 280 285

Ser Leu Asp Asp Arg Gly Ala Tyr Ser Leu Leu Ala Gln Leu Thr Met
290 295 300

Val Ala Glu Asp Val Ser Ala Gln Thr Gly Ile Pro Arg Ser Leu Ile
305 310 315 320

Ala Glu Ser Glu Leu Asn Asp Pro Lys Phe Val Thr Ser Arg Glu Ala
325 330 335

Gly Gly Phe Gly Leu Asp Ala Gln Trp Val Asp Asp Ile His His Ala
340 345 350

Leu His Ala Leu Val Ser Gly Glu Arg Asn Gly Tyr Tyr Ser Asp Phe
355 360 365

Gly Ser Val Asp Thr Leu Ala Lys Thr Leu Arg Glu Val Phe Glu His
370 375 380

Thr Gly Asn Tyr Ser Thr Tyr Arg Gly Arg Asn His Gly Arg Pro Val
385 390 395 400

His Pro Asp Ile Thr Pro Ala Ser Arg Phe Val Thr Tyr Thr Thr Thr
405 410 415

His Asp Gln Thr Gly Asn Arg Ala Ile Gly Asp Arg Pro Ser Thr Thr
420 425 430

Leu Thr Pro Glu Gln Gln Val Leu Lys Ala Ala Ile Ile Tyr Ser Ser
435 440 445

Pro Tyr Thr Pro Met Leu Phe Met Gly Glu Glu Phe Gly Ala Thr Thr
450 455 460

Pro Phe Ala Phe Phe Cys Ser His Thr Asp Pro Glu Leu Asn Arg Leu
465 470 475 480

Thr Ser Glu Gly Arg Lys Arg Glu Phe Ala Arg Leu Gly Trp Asn Ala
485 490 495

Asp Asp Ile Pro Ser Pro Glu Leu Glu Ser Thr Phe Thr Ser Ser Lys
500 505 510

Leu Asp Trp Glu Phe Thr Ala Glu Gln Arg Arg Ile Asn Asp Ala Tyr
515 520 525

Lys Gln Leu Leu His Leu Arg His Thr Leu Gly Phe Ser Gln Pro Asn
530 535 540

Leu Leu Thr Leu Glu Val Glu His Gly Glu Asn Trp Leu Ser Met Ala
545 550 555 560

Asn Gly Arg Gly Arg Ile Leu Ala Asn Phe Ser Asp Asp Thr Ile Thr
565 570 575

Val Pro Leu Gly Gly Glu Leu Ile Tyr Ser Phe Thr Ser Pro Thr Val
580 585 590

Thr Asp Thr Ser Thr Thr Leu Gln Pro Trp Gly Phe Ala Ile Leu Thr
595 600 605

Arg Asn
610

<210> 37
<211> 832

<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(832)
<223> RXN02355

<400> 37
atttttgacc ctccgggggt gatttaacct aaaattccac acaaacgtgt tcgaggtcat 60

tagattgata agcatctgtt gttaagaaag gtgacttcct atg tcc tcg att tcc 115
Met Ser Ser Ile Ser
1 5

cgc aag acc ggc gcg tca ctt gca gcc acc aca ctg ttg gca gcg atc 163
Arg Lys Thr Gly Ala Ser Leu Ala Ala Thr Thr Leu Leu Ala Ala Ile
10 15 20

gca ctg gcc ggt tgt agt tca gac tca agc tcc gac tcc aca gat tcc 211
Ala Leu Ala Gly Cys Ser Ser Asp Ser Ser Ser Asp Ser Thr Asp Ser
25 30 35

acc gct agc gaa ggc gca gac agc cgc ggc ccc atc acc ttt gcg atg 259
Thr Ala Ser Glu Gly Ala Asp Ser Arg Gly Pro Ile Thr Phe Ala Met
40 45 50

ggc aaa aac gac acc gac aaa gtc att ccg atc atc gac cgc tgg aac 307
Gly Lys Asn Asp Thr Asp Lys Val Ile Pro Ile Ile Asp Arg Trp Asn
55 60 65

gaa gcc cac ccc gat gag cag gta acg ctc aac gaa ctc gcc ggt gaa 355
Glu Ala His Pro Asp Glu Gln Val Thr Leu Asn Glu Leu Ala Gly Glu
70 75 80 85

gcc gac gcg cag cgc gaa acc ctc gtg caa tcc ctg cag gcc ggc aac 403
Ala Asp Ala Gln Arg Glu Thr Leu Val Gln Ser Leu Gln Ala Gly Asn
90 95 100

tct gac tac gac gtc atg gcg ctc gac gtc atc tgg acc gca gac ttc 451
Ser Asp Tyr Asp Val Met Ala Leu Asp Val Ile Trp Thr Ala Asp Phe
105 110 115

gcg gca aac caa tgg ctc gca cca ctt gaa ggc gac ctc gag gta gac 499
Ala Ala Asn Gln Trp Leu Ala Pro Leu Glu Gly Asp Leu Glu Val Asp
120 125 130

acc tcc gga ctg ctg caa tcc acc gtg gat tcc gca acc tac aac ggc 547
Thr Ser Gly Leu Leu Gln Ser Thr Val Asp Ser Ala Thr Tyr Asn Gly
135 140 145

acc ctc tac gca ctg cca cag aac acc aac ggc cag cta ctg ttc cgc 595
Thr Leu Tyr Ala Leu Pro Gln Asn Thr Asn Gly Gln Leu Leu Phe Arg
150 155 160 165

```
aac acc gaa atc atc cca gaa gca cca gca aac tgg gct gac ctc gtg   643
Asn Thr Glu Ile Ile Pro Glu Ala Pro Ala Asn Trp Ala Asp Leu Val
                170                 175                 180

gaa tcc tgc acg ctt gct gaa gaa gca ggc gtt gat tgc ctg acc act   691
Glu Ser Cys Thr Leu Ala Glu Glu Ala Gly Val Asp Cys Leu Thr Thr
            185                 190                 195

cag ctc aag cag tac gaa ggc ctt tca gtg aac acc atc ggc ttc atc   739
Gln Leu Lys Gln Tyr Glu Gly Leu Ser Val Asn Thr Ile Gly Phe Ile
        200                 205                 210

gaa ggt tgg gga ggc agc gtc cta gac gat gac ggc aaa cgt cac cgt   787
Glu Gly Trp Gly Gly Ser Val Leu Asp Asp Asp Gly Lys Arg His Arg
    215                 220                 225

aga cag cac gac ggc aag gca ggc ctt caa gcg ctt gtc gac ggc       832
Arg Gln His Asp Gly Lys Ala Gly Leu Gln Ala Leu Val Asp Gly
230                 235                 240
```

<210> 38
<211> 244
<212> PRT
<213> Corynebacterium glutamicum

<400> 38

```
Met Ser Ser Ile Ser Arg Lys Thr Gly Ala Ser Leu Ala Ala Thr Thr
  1               5                  10                  15

Leu Leu Ala Ala Ile Ala Leu Ala Gly Cys Ser Ser Asp Ser Ser Ser
             20                  25                  30

Asp Ser Thr Asp Ser Thr Ala Ser Glu Gly Ala Asp Ser Arg Gly Pro
         35                  40                  45

Ile Thr Phe Ala Met Gly Lys Asn Asp Thr Asp Lys Val Ile Pro Ile
     50                  55                  60

Ile Asp Arg Trp Asn Glu Ala His Pro Asp Glu Gln Val Thr Leu Asn
 65                  70                  75                  80

Glu Leu Ala Gly Glu Ala Asp Ala Gln Arg Glu Thr Leu Val Gln Ser
                 85                  90                  95

Leu Gln Ala Gly Asn Ser Asp Tyr Asp Val Met Ala Leu Asp Val Ile
            100                 105                 110

Trp Thr Ala Asp Phe Ala Ala Asn Gln Trp Leu Ala Pro Leu Glu Gly
        115                 120                 125

Asp Leu Glu Val Asp Thr Ser Gly Leu Leu Gln Ser Thr Val Asp Ser
    130                 135                 140

Ala Thr Tyr Asn Gly Thr Leu Tyr Ala Leu Pro Gln Asn Thr Asn Gly
145                 150                 155                 160
```

```
Gln Leu Leu Phe Arg Asn Thr Glu Ile Ile Pro Glu Ala Pro Ala Asn
                165                 170                 175

Trp Ala Asp Leu Val Glu Ser Cys Thr Leu Ala Glu Glu Ala Gly Val
            180                 185                 190

Asp Cys Leu Thr Thr Gln Leu Lys Gln Tyr Glu Gly Leu Ser Val Asn
        195                 200                 205

Thr Ile Gly Phe Ile Glu Gly Trp Gly Gly Ser Val Leu Asp Asp Asp
    210                 215                 220

Gly Lys Arg His Arg Arg Gln His Asp Gly Lys Ala Gly Leu Gln Ala
225                 230                 235                 240

Leu Val Asp Gly
```

<210> 39
<211> 609
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(586)
<223> RXN02909

<400> 39

```
caacgcgaat gaaaacgaac agcgagcagg tctataccca cgacgtcaac gtgtgggcta 60

atagtttcct ggattgtttg gcacagtcgg gagaaaactc atg aac cgc gca cga   115
                                            Met Asn Arg Ala Arg
                                              1               5

atc gcg acc ata ggc gtt ctt ccg ctt gct tta ctg ctg gcg tcc tgt   163
Ile Ala Thr Ile Gly Val Leu Pro Leu Ala Leu Leu Leu Ala Ser Cys
                10                  15                  20

ggt tca gac acc gtg gaa atg aca gat tcc acc tgg ttg gtg acc aat   211
Gly Ser Asp Thr Val Glu Met Thr Asp Ser Thr Trp Leu Val Thr Asn
            25                  30                  35

att tac acc gat cca gat gag tcg aat tcg atc agt aat ctt gtc att   259
Ile Tyr Thr Asp Pro Asp Glu Ser Asn Ser Ile Ser Asn Leu Val Ile
        40                  45                  50

tcc cag ccc agc tta gat ttt ggc aat tct tcc ctg tct ggt ttc act   307
Ser Gln Pro Ser Leu Asp Phe Gly Asn Ser Ser Leu Ser Gly Phe Thr
    55                  60                  65

ggc tgt gtg cct ttt acg ggg cgt gcg gaa ttc ttc caa aat ggt gag   355
Gly Cys Val Pro Phe Thr Gly Arg Ala Glu Phe Phe Gln Asn Gly Glu
70                  75                  80                  85
```

caa agc tct gtt ctg gat gcc gat tat gtg acc ttg tct tcc ctg gat 403
Gln Ser Ser Val Leu Asp Ala Asp Tyr Val Thr Leu Ser Ser Leu Asp
90 95 100

ttc gat aaa ctt ccc gat gat tgc caa gga caa gaa ctc aaa gtt cat 451
Phe Asp Lys Leu Pro Asp Asp Cys Gln Gly Gln Glu Leu Lys Val His
105 110 115

aac gag ctg gtt gat ctt ctg cct ggt tct ttt gaa atc tcc agg act 499
Asn Glu Leu Val Asp Leu Leu Pro Gly Ser Phe Glu Ile Ser Arg Thr
120 125 130

tct ggt tca gaa atc ttg ctg act agc gat gtc gat gaa ctc gat cgg 547
Ser Gly Ser Glu Ile Leu Leu Thr Ser Asp Val Asp Glu Leu Asp Arg
135 140 145

cca gca atc cgc ttg gtg tcc tgg atc gcg ccg aca tct taaggtgcca 596
Pro Ala Ile Arg Leu Val Ser Trp Ile Ala Pro Thr Ser
150 155 160

gggctttaaa gtg 609

<210> 40
<211> 162
<212> PRT
<213> Corynebacterium glutamicum

<400> 40
Met Asn Arg Ala Arg Ile Ala Thr Ile Gly Val Leu Pro Leu Ala Leu
1 5 10 15

Leu Leu Ala Ser Cys Gly Ser Asp Thr Val Glu Met Thr Asp Ser Thr
20 25 30

Trp Leu Val Thr Asn Ile Tyr Thr Asp Pro Asp Glu Ser Asn Ser Ile
35 40 45

Ser Asn Leu Val Ile Ser Gln Pro Ser Leu Asp Phe Gly Asn Ser Ser
50 55 60

Leu Ser Gly Phe Thr Gly Cys Val Pro Phe Thr Gly Arg Ala Glu Phe
65 70 75 80

Phe Gln Asn Gly Glu Gln Ser Ser Val Leu Asp Ala Asp Tyr Val Thr
85 90 95

Leu Ser Ser Leu Asp Phe Asp Lys Leu Pro Asp Asp Cys Gln Gly Gln
100 105 110

Glu Leu Lys Val His Asn Glu Leu Val Asp Leu Leu Pro Gly Ser Phe
115 120 125

Glu Ile Ser Arg Thr Ser Gly Ser Glu Ile Leu Leu Thr Ser Asp Val
130 135 140

```
Asp Glu Leu Asp Arg Pro Ala Ile Arg Leu Val Ser Trp Ile Ala Pro
145                 150                 155                 160

Thr Ser


<210> 41
<211> 1590
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(1567)
<223> RXS00349

<400> 41
tgtgtacatc acaatggaat tcggggctag agtatctggt gaaccgtgca taaacgacct 60

gtgattggac tctttttcct tgcaaaatgt tttccagcgg atg ttg agt ttt gcg   115
                                              Met Leu Ser Phe Ala
                                                1               5

acc ctt cgt ggc cgc att tca aca gtt gac gct gca aaa gcc gca cct   163
Thr Leu Arg Gly Arg Ile Ser Thr Val Asp Ala Ala Lys Ala Ala Pro
                10                  15                  20

ccg cca tcg cca cta gcc ccg att gat ctc act gac cat agt caa gtg   211
Pro Pro Ser Pro Leu Ala Pro Ile Asp Leu Thr Asp His Ser Gln Val
            25                  30                  35

gcc ggt gtg atg aat ttg gct gcg aga att ggc gat att ttg ctt tct   259
Ala Gly Val Met Asn Leu Ala Ala Arg Ile Gly Asp Ile Leu Leu Ser
        40                  45                  50

tca ggt acg tca aat agt gac acc aag gta caa gtt cga gca gtg acc   307
Ser Gly Thr Ser Asn Ser Asp Thr Lys Val Gln Val Arg Ala Val Thr
    55                  60                  65

tct gcg tac ggt ttg tac tac acg cac gtg gat atc acg ttg aat acg   355
Ser Ala Tyr Gly Leu Tyr Tyr Thr His Val Asp Ile Thr Leu Asn Thr
70                  75                  80                  85

atc acc atc ttc acc aac atc ggt gtg gag agg aag atg ccg gtc aac   403
Ile Thr Ile Phe Thr Asn Ile Gly Val Glu Arg Lys Met Pro Val Asn
                90                  95                  100

gtg ttt cat gtt gta ggc aag ttg gac acc aac ttc tcc aaa ctg tct   451
Val Phe His Val Val Gly Lys Leu Asp Thr Asn Phe Ser Lys Leu Ser
            105                 110                 115

gag gtt gac cgt ttg atc cgt tcc att cag gct ggt gcg acc ccg cct   499
Glu Val Asp Arg Leu Ile Arg Ser Ile Gln Ala Gly Ala Thr Pro Pro
        120                 125                 130
```

```
gag gtt gcc gag aaa atc ctg gac gag ttg gag caa tcc cct gcg tct   547
Glu Val Ala Glu Lys Ile Leu Asp Glu Leu Glu Gln Ser Pro Ala Ser
    135                 140                 145

tat ggt ttc cct gtt gcg ttg ctt ggc tgg gca atg atg ggt ggt gct   595
Tyr Gly Phe Pro Val Ala Leu Leu Gly Trp Ala Met Met Gly Gly Ala
150                 155                 160                 165

gtt gct gtg ctg ttg ggt ggt gga tgg cag gtt tcc cta att gct ttt   643
Val Ala Val Leu Leu Gly Gly Gly Trp Gln Val Ser Leu Ile Ala Phe
                170                 175                 180

att acc gcg ttc acg atc att gcc acg acg tca ttt ttg gga aag aag   691
Ile Thr Ala Phe Thr Ile Ile Ala Thr Thr Ser Phe Leu Gly Lys Lys
            185                 190                 195

ggt ttg cct act ttc ttc caa aat gtt gtt ggt ggt ttt att gcc acg   739
Gly Leu Pro Thr Phe Phe Gln Asn Val Val Gly Gly Phe Ile Ala Thr
        200                 205                 210

ctg cct gca tcg att gct tat tct ttg gcg ttg caa ttt ggt ctt gag   787
Leu Pro Ala Ser Ile Ala Tyr Ser Leu Ala Leu Gln Phe Gly Leu Glu
    215                 220                 225

atc aaa ccg agc cag atc atc gca tct gga att gtt gtg ctg ttg gca   835
Ile Lys Pro Ser Gln Ile Ile Ala Ser Gly Ile Val Val Leu Leu Ala
230                 235                 240                 245

ggt ttg aca ctc gtg caa tct ctg cag gac ggc atc acg ggc gct ccg   883
Gly Leu Thr Leu Val Gln Ser Leu Gln Asp Gly Ile Thr Gly Ala Pro
                250                 255                 260

gtg aca gca agt gca cga ttt ttc gaa aca ctc ctg ttt acc ggc ggc   931
Val Thr Ala Ser Ala Arg Phe Phe Glu Thr Leu Leu Phe Thr Gly Gly
            265                 270                 275

att gtt gct ggc gtg ggt ttg ggc att cag ctt tct gaa atc ttg cat   979
Ile Val Ala Gly Val Gly Leu Gly Ile Gln Leu Ser Glu Ile Leu His
        280                 285                 290

gtc atg ttg cct gcc atg gag tcc gct gca gca cct aat tat tcg tct   1027
Val Met Leu Pro Ala Met Glu Ser Ala Ala Ala Pro Asn Tyr Ser Ser
    295                 300                 305

aca ttc gcc cgc att atc gct ggt ggc gtc acc gca gcg gcc ttc gca   1075
Thr Phe Ala Arg Ile Ile Ala Gly Gly Val Thr Ala Ala Ala Phe Ala
310                 315                 320                 325

gtg ggt tgt tac gcg gag tgg tcc tcg gtg att att gcg ggg ctt act   1123
Val Gly Cys Tyr Ala Glu Trp Ser Ser Val Ile Ile Ala Gly Leu Thr
                330                 335                 340
```

```
gcg ctg atg ggt tct gcg ttt tat tac ctc ttc gtt gtt tat tta ggc   1171
Ala Leu Met Gly Ser Ala Phe Tyr Tyr Leu Phe Val Val Tyr Leu Gly
            345                 350                 355

ccc gtc tct gcc gct gcg att gct gca aca gca gtt ggt ttc act ggt   1219
Pro Val Ser Ala Ala Ala Ile Ala Ala Thr Ala Val Gly Phe Thr Gly
        360                 365                 370

ggt ttg ctt gcc cgt cga ttc ttg att cca ccg ttg att gtg gcg att   1267
Gly Leu Leu Ala Arg Arg Phe Leu Ile Pro Pro Leu Ile Val Ala Ile
    375                 380                 385

gcc ggc atc aca cca atg ctt cca ggt cta gca att tac cgc gga atg   1315
Ala Gly Ile Thr Pro Met Leu Pro Gly Leu Ala Ile Tyr Arg Gly Met
390                 395                 400                 405

tac gcc acc ctg aat gat caa aca ctc atg ggt ttc acc aac att gcg   1363
Tyr Ala Thr Leu Asn Asp Gln Thr Leu Met Gly Phe Thr Asn Ile Ala
                410                 415                 420

gtt gct tta gcc act gct tca tca ctt gcc gct ggc gtg gtt ttg ggt   1411
Val Ala Leu Ala Thr Ala Ser Ser Leu Ala Ala Gly Val Val Leu Gly
            425                 430                 435

gag tgg att gcc cgc agg cta cgt cgt cca cca cgc ttc aac cca tac   1459
Glu Trp Ile Ala Arg Arg Leu Arg Arg Pro Pro Arg Phe Asn Pro Tyr
        440                 445                 450

cgt gca ttt acc aag gcg aat gag ttc tcc ttc cag gag gaa gct gag   1507
Arg Ala Phe Thr Lys Ala Asn Glu Phe Ser Phe Gln Glu Glu Ala Glu
    455                 460                 465

cag aat cag cgc cgg cag aga aaa cgt cca aag act aat cag aga ttc   1555
Gln Asn Gln Arg Arg Gln Arg Lys Arg Pro Lys Thr Asn Gln Arg Phe
470                 475                 480                 485

ggt aat aaa agg taaaaatcaa cctgcttagg cgt                        1590
Gly Asn Lys Arg
```

<210> 42
<211> 489
<212> PRT
<213> Corynebacterium glutamicum

<400> 42

```
Met Leu Ser Phe Ala Thr Leu Arg Gly Arg Ile Ser Thr Val Asp Ala
  1               5                  10                  15

Ala Lys Ala Ala Pro Pro Pro Ser Pro Leu Ala Pro Ile Asp Leu Thr
             20                  25                  30

Asp His Ser Gln Val Ala Gly Val Met Asn Leu Ala Ala Arg Ile Gly
         35                  40                  45
```

```
Asp Ile Leu Leu Ser Ser Gly Thr Ser Asn Ser Asp Thr Lys Val Gln
     50                  55                  60

Val Arg Ala Val Thr Ser Ala Tyr Gly Leu Tyr Tyr Thr His Val Asp
 65                  70                  75                  80

Ile Thr Leu Asn Thr Ile Thr Ile Phe Thr Asn Ile Gly Val Glu Arg
                 85                  90                  95

Lys Met Pro Val Asn Val Phe His Val Val Gly Lys Leu Asp Thr Asn
            100                 105                 110

Phe Ser Lys Leu Ser Glu Val Asp Arg Leu Ile Arg Ser Ile Gln Ala
        115                 120                 125

Gly Ala Thr Pro Pro Glu Val Ala Glu Lys Ile Leu Asp Glu Leu Glu
    130                 135                 140

Gln Ser Pro Ala Ser Tyr Gly Phe Pro Val Ala Leu Leu Gly Trp Ala
145                 150                 155                 160

Met Met Gly Gly Ala Val Ala Val Leu Leu Gly Gly Gly Trp Gln Val
                165                 170                 175

Ser Leu Ile Ala Phe Ile Thr Ala Phe Thr Ile Ile Ala Thr Thr Ser
            180                 185                 190

Phe Leu Gly Lys Lys Gly Leu Pro Thr Phe Phe Gln Asn Val Val Gly
        195                 200                 205

Gly Phe Ile Ala Thr Leu Pro Ala Ser Ile Ala Tyr Ser Leu Ala Leu
    210                 215                 220

Gln Phe Gly Leu Glu Ile Lys Pro Ser Gln Ile Ile Ala Ser Gly Ile
225                 230                 235                 240

Val Val Leu Leu Ala Gly Leu Thr Leu Val Gln Ser Leu Gln Asp Gly
                245                 250                 255

Ile Thr Gly Ala Pro Val Thr Ala Ser Ala Arg Phe Phe Glu Thr Leu
            260                 265                 270

Leu Phe Thr Gly Gly Ile Val Ala Gly Val Gly Leu Gly Ile Gln Leu
        275                 280                 285

Ser Glu Ile Leu His Val Met Leu Pro Ala Met Glu Ser Ala Ala Ala
    290                 295                 300

Pro Asn Tyr Ser Ser Thr Phe Ala Arg Ile Ile Ala Gly Gly Val Thr
305                 310                 315                 320

Ala Ala Ala Phe Ala Val Gly Cys Tyr Ala Glu Trp Ser Ser Val Ile
                325                 330                 335

Ile Ala Gly Leu Thr Ala Leu Met Gly Ser Ala Phe Tyr Tyr Leu Phe
            340                 345                 350
```

```
Val Val Tyr Leu Gly Pro Val Ser Ala Ala Ala Ile Ala Ala Thr Ala
        355                 360                 365

Val Gly Phe Thr Gly Gly Leu Leu Ala Arg Arg Phe Leu Ile Pro Pro
    370                 375                 380

Leu Ile Val Ala Ile Ala Gly Ile Thr Pro Met Leu Pro Gly Leu Ala
385                 390                 395                 400

Ile Tyr Arg Gly Met Tyr Ala Thr Leu Asn Asp Gln Thr Leu Met Gly
                405                 410                 415

Phe Thr Asn Ile Ala Val Ala Leu Ala Thr Ala Ser Ser Leu Ala Ala
            420                 425                 430

Gly Val Val Leu Gly Glu Trp Ile Ala Arg Arg Leu Arg Arg Pro Pro
        435                 440                 445

Arg Phe Asn Pro Tyr Arg Ala Phe Thr Lys Ala Asn Glu Phe Ser Phe
    450                 455                 460

Gln Glu Glu Ala Glu Gln Asn Gln Arg Arg Gln Arg Lys Arg Pro Lys
465                 470                 475                 480

Thr Asn Gln Arg Phe Gly Asn Lys Arg
                485
```

<210> 43
<211> 440
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (1)..(417)
<223> RXS03183

<400> 43

```
gaa gcc gaa gca acc gca ggc aaa ttc gaa gta cag ccc ctc gta ggt   48
Glu Ala Glu Ala Thr Ala Gly Lys Phe Glu Val Gln Pro Leu Val Gly
  1               5                  10                  15

aaa gac ggc gtc ggc gta tcc acc ctt ggt ggc tac aac aac ggc atc   96
Lys Asp Gly Val Gly Val Ser Thr Leu Gly Gly Tyr Asn Asn Gly Ile
             20                  25                  30

aac gtc aac tcc gaa aac aag gca acc gcc cgc gac ttc atc gaa ttc   144
Asn Val Asn Ser Glu Asn Lys Ala Thr Ala Arg Asp Phe Ile Glu Phe
         35                  40                  45

atc atc aac gaa gag aac caa acc tgg ttc gcg gac aac tcc ttc cca   192
Ile Ile Asn Glu Glu Asn Gln Thr Trp Phe Ala Asp Asn Ser Phe Pro
     50                  55                  60
```

```
cca gtt ctg gca tcc atc tac gat gat gag tcc ctt gtt gag cag tac   240
Pro Val Leu Ala Ser Ile Tyr Asp Asp Glu Ser Leu Val Glu Gln Tyr
 65                  70                  75                  80

cca tac ctg cca gca ctg aag gaa tcc ctg gaa aac gca gca cca cgc   288
Pro Tyr Leu Pro Ala Leu Lys Glu Ser Leu Glu Asn Ala Ala Pro Arg
                85                  90                  95

cca gtg tct cct ttc tac cca gcc atc tcc aag gca atc cag gac aac   336
Pro Val Ser Pro Phe Tyr Pro Ala Ile Ser Lys Ala Ile Gln Asp Asn
            100                 105                 110

gcc tac gca gcg ctt aac ggc aac gtc gac gtt gac cag gca acc acc   384
Ala Tyr Ala Ala Leu Asn Gly Asn Val Asp Val Asp Gln Ala Thr Thr
        115                 120                 125

gat atg aag gca gcg atc gaa aac gct tcc agc tagttcggta atttagttca 437
Asp Met Lys Ala Ala Ile Glu Asn Ala Ser Ser
    130                 135

ttc                                                               440
```

<210> 44
<211> 139
<212> PRT
<213> Corynebacterium glutamicum

<400> 44

```
Glu Ala Glu Ala Thr Ala Gly Lys Phe Glu Val Gln Pro Leu Val Gly
  1               5                  10                  15

Lys Asp Gly Val Gly Val Ser Thr Leu Gly Gly Tyr Asn Asn Gly Ile
             20                  25                  30

Asn Val Asn Ser Glu Asn Lys Ala Thr Ala Arg Asp Phe Ile Glu Phe
         35                  40                  45

Ile Ile Asn Glu Glu Asn Gln Thr Trp Phe Ala Asp Asn Ser Phe Pro
     50                  55                  60

Pro Val Leu Ala Ser Ile Tyr Asp Asp Glu Ser Leu Val Glu Gln Tyr
 65                  70                  75                  80

Pro Tyr Leu Pro Ala Leu Lys Glu Ser Leu Glu Asn Ala Ala Pro Arg
                85                  90                  95

Pro Val Ser Pro Phe Tyr Pro Ala Ile Ser Lys Ala Ile Gln Asp Asn
            100                 105                 110

Ala Tyr Ala Ala Leu Asn Gly Asn Val Asp Val Asp Gln Ala Thr Thr
        115                 120                 125

Asp Met Lys Ala Ala Ile Glu Asn Ala Ser Ser
    130                 135
```

<210> 45
<211> 1212
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (101)..(1189)
<223> RXC00874

<400> 45
agctgttccc taccattgct gaacgggagt ggattgtcac tttagcccct cacggattct 60

tctggtttga tctcaccgcc gatgaaaagg acgatatgga atg agc att ggc caa 115
Met Ser Ile Gly Gln
1 5

cac atc atc acc gag cgt ttc tac ggc gcc aag tcc cac acc atc gac 163
His Ile Ile Thr Glu Arg Phe Tyr Gly Ala Lys Ser His Thr Ile Asp
10 15 20

aac gta gat att gtg ttg tcc cgc gaa tgt ggc gag aac act ttg gct 211
Asn Val Asp Ile Val Leu Ser Arg Glu Cys Gly Glu Asn Thr Leu Ala
25 30 35

gta gtg cgc atc aac aat gcg ctg tat cag ttg ttg gtc aat gat gat 259
Val Val Arg Ile Asn Asn Ala Leu Tyr Gln Leu Leu Val Asn Asp Asp
40 45 50

ggc aaa gat gtt ctc aac gac cac gta gaa gag gtc ggt gcg agt ttc 307
Gly Lys Asp Val Leu Asn Asp His Val Glu Glu Val Gly Ala Ser Phe
55 60 65

gga gca tgg act ggc agc tct gct ttt ccc att ggc cct ttc act cca 355
Gly Ala Trp Thr Gly Ser Ser Ala Phe Pro Ile Gly Pro Phe Thr Pro
70 75 80 85

ctc ggc aca gaa caa tcc aat agc tct ttc atc acc gcc gac aat aaa 403
Leu Gly Thr Glu Gln Ser Asn Ser Ser Phe Ile Thr Ala Asp Asn Lys
90 95 100

gcg atc gtg aaa tac ttc cgc aaa tta gaa tcc ggg caa aac ccc gat 451
Ala Ile Val Lys Tyr Phe Arg Lys Leu Glu Ser Gly Gln Asn Pro Asp
105 110 115

gtg gag cta att tct aaa att tcc tcc tgc ccc aac atc gcg ccc atc 499
Val Glu Leu Ile Ser Lys Ile Ser Ser Cys Pro Asn Ile Ala Pro Ile
120 125 130

ctg ggt ttt tcc tcc gct gag atc tcc ggg gct aac tac acc ctg gtc 547
Leu Gly Phe Ser Ser Ala Glu Ile Ser Gly Ala Asn Tyr Thr Leu Val
135 140 145

```
atg gcg cag cag tac gtt cca ggt ttg gat ggc tgg tca cac gcg ctg   595
Met Ala Gln Gln Tyr Val Pro Gly Leu Asp Gly Trp Ser His Ala Leu
150                 155                 160                 165

act act acc tct ggc agc ttt gca gag gat gca gaa aag atc ggc gaa   643
Thr Thr Thr Ser Gly Ser Phe Ala Glu Asp Ala Glu Lys Ile Gly Glu
                170                 175                 180

gcc acc cgc aat gtt cac act gct ctt gca tcg gcc ttc cct act cgg   691
Ala Thr Arg Asn Val His Thr Ala Leu Ala Ser Ala Phe Pro Thr Arg
            185                 190                 195

gta gtt ccc gta gaa gca ctc gcc gat gcg ctc act acc cgc ctt aat   739
Val Val Pro Val Glu Ala Leu Ala Asp Ala Leu Thr Thr Arg Leu Asn
        200                 205                 210

gaa cta atc tcc caa gca ccc gaa atc gcc cgc ttc aaa gaa gca gcc   787
Glu Leu Ile Ser Gln Ala Pro Glu Ile Ala Arg Phe Lys Glu Ala Ala
    215                 220                 225

atc gac ctc tac caa tcg ttg gaa ggc gaa gcc cac atc caa cgc atc   835
Ile Asp Leu Tyr Gln Ser Leu Glu Gly Glu Ala His Ile Gln Arg Ile
230                 235                 240                 245

cac ggt gac ctc cac ttg ggg cag ctc atc aaa acc ccc gaa cgc tac   883
His Gly Asp Leu His Leu Gly Gln Leu Ile Lys Thr Pro Glu Arg Tyr
                250                 255                 260

atc ctc atc gat ttc gaa ggc gaa cct gcc cgc cca ctt aat caa cga   931
Ile Leu Ile Asp Phe Glu Gly Glu Pro Ala Arg Pro Leu Asn Gln Arg
            265                 270                 275

cgc ctc ccc gac tct ccc ctg aaa gat ctc gcc ggc atc atc aga tcc   979
Arg Leu Pro Asp Ser Pro Leu Lys Asp Leu Ala Gly Ile Ile Arg Ser
        280                 285                 290

atc gac tac gca gcc tac ttc gac ggc gaa cac acc caa tgg gcc aac   1027
Ile Asp Tyr Ala Ala Tyr Phe Asp Gly Glu His Thr Gln Trp Ala Asn
    295                 300                 305

gaa gcc acc gcg cta ttc ctc gac ggc tac gga tca att gaa gac caa   1075
Glu Ala Thr Ala Leu Phe Leu Asp Gly Tyr Gly Ser Ile Glu Asp Gln
310                 315                 320                 325

gaa ctc ctc aat gcc tac att ctg gac aag gcg ttg tac gag gtt gcc   1123
Glu Leu Leu Asn Ala Tyr Ile Leu Asp Lys Ala Leu Tyr Glu Val Ala
                330                 335                 340

tat gaa ata aac aac cgc ccc gac tgg gtg aaa atc cca ctc gag gcg   1171
Tyr Glu Ile Asn Asn Arg Pro Asp Trp Val Lys Ile Pro Leu Glu Ala
            345                 350                 355
```

gtc gaa agg ctt cta gac tagttagtta ctctgcgtca aac 1212
Val Glu Arg Leu Leu Asp
360

<210> 46
<211> 363
<212> PRT
<213> Corynebacterium glutamicum

<400> 46
Met Ser Ile Gly Gln His Ile Ile Thr Glu Arg Phe Tyr Gly Ala Lys
1 5 10 15

Ser His Thr Ile Asp Asn Val Asp Ile Val Leu Ser Arg Glu Cys Gly
20 25 30

Glu Asn Thr Leu Ala Val Val Arg Ile Asn Asn Ala Leu Tyr Gln Leu
35 40 45

Leu Val Asn Asp Asp Gly Lys Asp Val Leu Asn Asp His Val Glu Glu
50 55 60

Val Gly Ala Ser Phe Gly Ala Trp Thr Gly Ser Ser Ala Phe Pro Ile
65 70 75 80

Gly Pro Phe Thr Pro Leu Gly Thr Glu Gln Ser Asn Ser Ser Phe Ile
85 90 95

Thr Ala Asp Asn Lys Ala Ile Val Lys Tyr Phe Arg Lys Leu Glu Ser
100 105 110

Gly Gln Asn Pro Asp Val Glu Leu Ile Ser Lys Ile Ser Ser Cys Pro
115 120 125

Asn Ile Ala Pro Ile Leu Gly Phe Ser Ser Ala Glu Ile Ser Gly Ala
130 135 140

Asn Tyr Thr Leu Val Met Ala Gln Gln Tyr Val Pro Gly Leu Asp Gly
145 150 155 160

Trp Ser His Ala Leu Thr Thr Thr Ser Gly Ser Phe Ala Glu Asp Ala
165 170 175

Glu Lys Ile Gly Glu Ala Thr Arg Asn Val His Thr Ala Leu Ala Ser
180 185 190

Ala Phe Pro Thr Arg Val Val Pro Val Glu Ala Leu Ala Asp Ala Leu
195 200 205

Thr Thr Arg Leu Asn Glu Leu Ile Ser Gln Ala Pro Glu Ile Ala Arg
210 215 220

Phe Lys Glu Ala Ala Ile Asp Leu Tyr Gln Ser Leu Glu Gly Glu Ala
225 230 235 240

His Ile Gln Arg Ile His Gly Asp Leu His Leu Gly Gln Leu Ile Lys
245 250 255

Thr Pro Glu Arg Tyr Ile Leu Ile Asp Phe Glu Gly Glu Pro Ala Arg
260 265 270

Pro Leu Asn Gln Arg Arg Leu Pro Asp Ser Pro Leu Lys Asp Leu Ala
275 280 285

Gly Ile Ile Arg Ser Ile Asp Tyr Ala Ala Tyr Phe Asp Gly Glu His
290 295 300

Thr Gln Trp Ala Asn Glu Ala Thr Ala Leu Phe Leu Asp Gly Tyr Gly
305 310 315 320

Ser Ile Glu Asp Gln Glu Leu Leu Asn Ala Tyr Ile Leu Asp Lys Ala
325 330 335

Leu Tyr Glu Val Ala Tyr Glu Ile Asn Asn Arg Pro Asp Trp Val Lys
340 345 350

Ile Pro Leu Glu Ala Val Glu Arg Leu Leu Asp
355 360

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 00/13143

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/54 C12N9/10 C12N1/20 C12N15/74 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE SWISS-PROT [Online] EBI; 15 July 1998 (1998-07-15) "5-Methyltetrahydrofolate-homocysteine methyltransferase (EC 2.1.1.13) of Mycobacterium tuberculosis" XP002175756 Acc. No. O33259 ---	3-13, 15-17, 19-25, 28-33,36
X	DATABASE EMBL/GENBANK/DDBJ [Online] EBI; 10 July 1997 (1997-07-10) "Mycobacterium tuberculosis H37Rv complete genome; segment 95-162" XP002175757 Acc. No. Z97559 --- -/--	3-13, 15-17, 23-25, 28-33,36

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 August 2001	14 11. 2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Mata-Vicente, M

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 00/13143

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	FUJII, K. ET AL.: "Vitamin B12-dependent Methionine Biosynthesis and Its Metabolic Role in Corynebacterium simplex ATCC 6946, a Vitamin B12-producing and Hydrocarbon-utilizable Bacterium" AGR. BIOL. CHEM., vol. 36, no. 13, 1972, pages 2323-2334, XP001016120 page 2332, left-hand column, paragraph 2 ---	1-37
A	GROSSMANN, K. ET AL: "Rapid Cloning of metK encoding methionine adenosyltransferase from Corynebacterium glutamicum by screeening a genomic library on a high density colony-array" FEMS MICROBIOLOGY LETTERS, AMSTERDAM, NL, vol. 193, no. 1, 1 December 2000 (2000-12-01), pages 99-103, XP000984551 ISSN: 0378-1097 abstract ---	1-37
A	PARK, S-D ET AL: "Isolation and Analysis of metA, a Methionine Biosynthetic Gene Encoding Homoserine Acetyltransferase in Corynebacterium glutamicum" MOLECULAR AND CELLS, KOREAN SOCIETY FOR MOLECULAR SOCIETY, KR, vol. 8, no. 3, 30 June 1998 (1998-06-30), pages 286-294, XP001002218 abstract -----	1-37

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 00/13143

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

As far as an "in vivo" method is concerned, claim 34 is directed to a diagnostic method practised on the human/animal body and the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-37 (all partially)

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-37 (partially)

An isolated Corynebacterium glutamicum nucleic acid molecule consisting of SEQ ID NO:1, fragments and variants thereof and polynucleotides capable of hybridizing with them; vectors containing said nucleic acid molecules, host cells containing the vectors and method to produce a polypeptide comprising culturing said host cells.
An isolated polypeptide comprising SEQ ID NO:2, fragments and variants thereof.
Methods for producing fine chemicals comprising the use of any of the above mentioned molecules or cells.
Method for diagnosing the presence or activity of Corynebacterium diphteriae comprising detecting the presence of SEQ ID NOs:1 or 2.

2. Claims: 1-37 (partially)

Idem as in subject 1, but restricted to SEQ ID NOs:3 and 4.